ROBUST DUAL MEMORY CONTROL CHARTING SCHEMES

BY

RAJI, ISHAQ ADEYANJU

A Thesis Presented to the
DEANSHIP OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

In

APPLIED STATISTICS

December 2015

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN- 31261, SAUDI ARABIA

DEANSHIP OF GRADUATE STUDIES

This thesis, written by RAJI, ISHAQ ADEYANJU under the direction his thesis advisor and approved by his thesis committee, has been presented and accepted by the Dean of Graduate Studies, in partial fulfillment of the requirements for the degree of **MASTER OF SCIENCE IN APPLIED STATISTICS.**

Dr. Muhammad Riaz
(Advisor)

Dr. Husain Salem Al-Attas
Department Chairman

Dr. Nasir Abbas
(Co-Advisor)

Dr. Salam A. Zummo
Dean of Graduate Studies

Dr. Saddam Akbar Abbasi
(Member)

3/1/16
Date

Dr. Walid Sabah Al-Sabah
(Member)

Dr. Mohammad H Omar
(Member)

To my Parents,

Siblings,

Wife

And my Son Jalwaan

ACKNOWLEDGMENTS

All adorations are due to Almighty Allah, who has taught mankind by pen, He taught them what they knew not. Surely, with Him all things are easy and possible, and without Him never. My gratitude is for Him, from the beginning, this moment and till eternity.

I'm totally indebted to my parents, for their care, prayers and supports. May Allah save and spare their lives to reap the fruit of their labor, for their labor is so tremendous. "O' Allah, have mercy on them (both) as they did to me when I was little"

My undiluted appreciation goes to my Supervisor, Co-Advisor and other members of my thesis committee for their quality supervision, constructive criticism and propelling support throughout my learning period and especially the research time. I must use this medium, to also appreciate Prof. Anwar Joarder, a history of probability, you're always remembered for your impacted knowledge is never forgotten.

To my dear wife, Ummu Jalwaan, thanks for the love, care, courage, support and confidence and most importantly for taking care of my son. Jazakillahu khairan.

To my siblings, family members, friends home and abroad, and to everyone who has directly or indirectly contributed to the actualization of this program. Thank you very much. I say Jazakumullahu khairan to you all.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ARL: Average Run Length

ARL_0: In- control Average Run Length

ARL_1: Out-of-control Average Run Length

CL: Center Line

CUSUM: Cumulative Sum

DCUSUM: Dual Cumulative Sum

DEWMA: Double Exponentially Weighted Moving Average

EWMA: Exponentially Weighted Moving Average

FIR: Fast Initial Response

HL: Hodges-Lehmann estimator

LCNE: Location-Contaminated Normal Environment

LCL: Lower Control Limit

MCE: Mixed CUSUM-EWMA

MEDC: Mixed EWMA-Dual-CUSUM

MEC: Mixed EWMA-CUSUM

MR: Midrange estimator

SDRL: Standard Deviation of Run Length

SPC: Statistical Process Control

TM: Trimean estimator

UCL: Upper Control Limit

VCNE: Variance Contaminated Normal Environment

ABSTRACT

Full Name : RAJI, ISHAQ ADEYANJU

Thesis Title : ROBUST DUAL MEMORY CONTROL CHARTING SCHEMES

Major Field : APPLIED STATISTICS

Date of Degree : DECEMBER 2015

Control Chart is an expedient tool of the statistical process control amongst others. Of larger importance is the memory type of control charts -due to its sensitivity and quick ability to detect small and moderate shifts- which are basically, Cumulative Sum (CUSUM) and Exponentially Weighted Moving Average (EWMA). These charts are very effective and sensitive to small and moderate shifts, unlike the memory-less type. A modification of these two charts is extending them to double charts, by repeating the structure, and monitor them simultaneously in order to increase their efficiency. In this research study, we proposed some robust dual memory control charting schemes and explored them with robust location estimators.

Dual CUSUM, Double EWMA and Mixed EWMA Dual-CUSUM control charting schemes were the proposed schemes explored with some robust estimators. We evaluated their performances based on their ARL's values under contaminated and uncontaminated normal environments. We made a brief comparison of these charts and buttressed the claims with their application on real life data set.

ARABIC ABSTRACT

ملخص الرسالة

الاسم الكامل: إسحاق أديي انجو راجى

عنوان الرسالة: برامج مراقبة الرسوم اليبانية الذاكرة المزدوجة القوية

التخصص: الإحصاء

تاريخ الدرجة العلمية: دسيمبر 2015

مراقبة الرسم البياني أداة نفعية لمراقبة العملية الإحصائية من غيرها من الآدات. والرسوم اليبانية ذات الذاكرة هي عليا أهمية – نظرا لحساسيتها وقدرتها السريعة لكشف التحولات الصغيرة والمتوسطة– التى هي فى الأساس: مجموع التراكمي والمتوسط المتحرك الموزون تصاعديا. وهذه المخططات هي فعالة جدا وحساسة للتغيرات الصغيرة والمتوسطة، على عكس نوع غير– الذاكرة ومن تعديل هذه المخططات امتداده إلى اثنتين، بتكرار هيكلها ومراقبتها في وقت واحد من أجل زيادة كفاءتها. في هذه الدراسة البحثية، بعض المخططات البيانية مراقبة ذاكرة مزدوجة قوية واستكشاف هذه المخططات بالمقدرات الموقعية القوية

CUSUM مزدوجEWMA مزدوج و EWMA Dual-CUSUMالمختلطة هي المخططات المقترحة،

وتم استكشافها مع بعض المقدرات القوية.قمنا ARL الخاصة بها في بيئات طبيعية الملوثة وغير الملوثة بتقييم أدائها على وتم إجراء مقارنة موجزة عن هذه المخططات والالتفاف على المطالبات مع التطبيق على مجموعة البيانات الحياة الحقيقية

CHAPTER 1

1. INTRODUCTION

Statistical Process Control (SPC) is a set of techniques used to control and monitor the special causes of variation in manufacturing or service processes, since the natural variations are inevitably part of the processes. These techniques are called the seven (7) tools of SPC, namely: Cause and Effect diagram, Pareto charts, histogram, flow charts, probability plots, scatter diagram and Control chart [1]. It is truism that amongst these techniques, Control Chart is the most efficient and widely use of all. Processes are statistically, said to be in an in-control state even with the presence of natural variations, but they are in an out-of-control state due to the existence of un-natural variations depending on the magnitude of the variations and the corresponding control charts. Based on this, control charts can be classified with respect to the magnitude of variation they are designed to control and measure which are the memory-less charts and the memory type. Another categorization of control charts is based on the number quality characteristics they measure, if it is one, univariate, and if more than two, its multivariate. These two types of classification can be carried out to measure both location or/and spread parameters for the process under consideration.

1.1 The Memory-Less Control Chart

The memory-less control chart uses the present information of the process, truncating the past information in constructing the chart statistic, which makes it less sensitive to detect small and moderate shift, but highly sensitive to large shifts in the process. An example of

the memory-less chart is the Shewhart chart named after Shewhart [2] which has been extended and modified to multivariate set up.

1.2 The Memory-Type Control Chart

As the name sound, this type of chart keeps into memory, the past event of the process. Unlike the memory-less chart, it makes use of both previous and past information of the process in constructing the chart statistics and by so doing makes it more sensitive to small and moderate shifts in the process under consideration. The two major examples for this type are Cumulative Sum (CUSUM) by Page [3] and Exponential Weighted Moving Average (EWMA) by Robert [4].

1.3 Modifications of the Memory Control Charts

The importance and effectiveness of the memory charts have attracted the interest of many researchers, leading to many modifications and improvements in the field. In the progression of increasing the efficiency and performance of these charts, various type of charts have being in the literature, some of which include: Shewhart-CUSUM chart by Lucas [5], he combined a memory and memory-less type of chart, such that the Shewhart statistic deals with large shifts, and the CUSUM statistic takes into account, the small and moderate shifts. Shewhart-EWMA chart was proposed by Lucas and Saccucci [6] , to control both large and small shifts with a single chart. Lucas and Crosier [7] introduced a Fast Initial Response (FIR) to the classical CUSUM in order to detect small and moderate shifts faster than the ordinary classical CUSUM. Abbas et al [8] also designed a mixed EWMA-CUSUM chart for enhancing the sensitivity of the chart to smaller magnitude. Other modifications in the literature are the omnibus EWMA method by Domangue &

Patch [9], adaptive EWMA chart by Capizi and Masarotto [10], also Borror et al[11], extensively investigated performance of EWMA chart under non-normal distributions.

Therefore, this thesis work aims to improve on the sensitivity of the memory charts – like other scholars in literature did- by extending the basic classical memory charts to dual/double and exploring these schemes in contaminated normal environments with some estimators to evaluate their robustness. The research also spreads its feather to proposing a new mixture of the dual-memory chart with a classical memory chart and exploring with same set of estimators in uncontaminated and contaminated normal environments to evaluate their robustness. Finally these set of schemes were compared and were applied in real-life situation to illustrate their applicability

CHAPTER 2

ROBUST DUAL-CUSUM CONTROL CHARTS FOR CONTAMINATED PROCESSES

2.1 INTRODUCTION

The trending modifications on the memory type charts (EWMA and CUSUM) are so numerous. On the motivation of enhancing the sensitivity of the CUSUM chart, Zhao et al [12] designed a Dual CUSUM (DCUSUM) scheme that combines a two one-sided CUSUM statistics and monitor them simultaneously to detect a range of small shifts in a process. This idea was an adjustment and improvement to that pioneered by Lordens [13], where he combined a set of infinite numbers of classical CUSUM statistics, and monitor them simultaneously to also detect a range of small shifts in a process for univariate set-up. He proved the asymptotical optimality feature of this chart as having the smallest out-of-control Average Run Length (ARL) when the in-control ARL tends to infinity. Unfortunately, there was no empirical or simulation illustration in literature to buttress his claim, due to the difficulty that arises in its computation. The simile of this is sweeping a piece of land in an infinite number of time, at a point in time, it becomes unnecessary and a complete waste of time and effort, while a definite number of time will be economically and efficiently reasonable.

The underlying assumptions of all the aforementioned charts are that the environment of these processes are normal, uncontaminated and free from outliers. In a situation where any of these assumptions is not met, then the efficiency of the corresponding chart becomes questionable. One of the ways to address this matter is to use some robust charts that can

not only resist the violation of the assumptions but also maintain the efficiency of their performances, whether the charts are measuring location or dispersion parameters. Series of publications are available in the literature, that incorporate robustness in the charts, most of whose techniques are based on transforming the data into non-parametric statistics, then adopting the CUSUM scheme with the transformed statistics. To mention a few, a robust control chart for estimating dispersion was designed by Tatum in 1997 [14], Mokhtar and Moustafa [15] constructed another robust control chart using the Hodges-Lehman estimator for location parameter and Shamos-Bickel-Lehman estimator for dispersion parameter. Lately, Riaz [16] had a comprehensive study on control charts whose ideal assumptions has been violated, their properties and effects on their performances were discussed in details. Schoonhoven et al [17] also have a similar publication in 2011.

On the contrary, L. Yang et al [18], proposed a robust CUSUM control chart using the sample median as the chart statistic, whom observed data were not transformed to a non-parametric statistic. In extension, Nazir et al [19] designed a set of robust CUSUM control chart with some robust location estimators and compared their efficiencies and robustness for phase II samples. In this article, we proposed a set of robust Dual CUSUM control charting schemes with five different location estimators (mean, median, midrange, trimean and Hodges-Lehmann) for monitoring the phase II samples of a process, keeping the mean estimator as a yardstick for measuring their robustness. We evaluated these charts based on their run length distribution (ARL and SDRL) and their performances were examined in different environments, the uncontaminated normal environment as well as the location and dispersion contaminated environment.

The remainder of this chapter is organized as follows: in the next section, we discuss the classical CUSUM control chart versus the construction and design of the Dual CUSUM (DCUSUM) control chart. Section 3 is a detailed description of the proposed estimators and their corresponding DCUSUM charts. While the performance evaluation of these proposed charts are presented in section 4, section 5 buttressed these with application on a real life dataset. Finally, summary and conclusion appears in section 6.

2.2 CLASSICAL CUSUM & DUAL CUSUM CONTROL CHARTS

In this section, we describe in details the Cumulative Sum (CUSUM) chart and its counterpart, the dual-CUSUM (DCUSUM) charts for process monitoring

2.2.1 The Classical CUSUM Chart

The classical CUSUM chart pioneered by Page[3], is now one of the most popular commonly used control chart for monitoring small and moderate shifts in processes. This chart, like other counterparts, can be used for individual data and subgroups (i.e. sample size n > 1). The classical CUSUM control chart comprises of two statistics $C_i^+ \; and \; C_i^-$ which are plotted against a single limit interval H. the statistics are defined as:

$$\left.\begin{aligned} C_i^+ &= \max\left[0, \left(\hat{\theta}_i - \mu_0\right) - K + C_{i-1}^+\right] \\ C_i^- &= \max\left[0, -\left(\hat{\theta}_i - \mu_0\right) - K + C_{i-1}^-\right] \end{aligned}\right\}$$

Where i is the subgroup number, $\hat{\theta}_i$ is the estimator of the study variable under consideration. μ_0 is the target mean of the study variable, $C_0^+ = C_0^- = 0$ and $K = k\sigma_{\bar{X}}$ is the reference value of the CUSUM scheme. The next task is to plot each of these two statistics against $H = h\sigma_{\bar{X}}$, if $C_i^+ > H$ for any value of i we conclude that the process has

shifted upward, and if $C_i^- > H$ at any point i, the process mean has moved downwards. When n = 1, $\hat{\theta}_i$ is replaced by the corresponding individual observation X_i. The parameters that define this chart are K and H, which are to be selected to a predetermined average run length of interest. Therefore, a careful selection of these parameters should be put in place in order not to affect the performance of the chart. Researchers have published a number of articles where the selection of these parameters and its effect on the chart were studied in details. The method explained above, is the tabular method of evaluating the CUSUM chart; the preferred and most common in use. The second method of CUSUM chart evaluation is the V-mask method. This uncommon method plots the normalized deviations of the mean. Its computational ambiguity and complexity made it loses research improvements and practical adoptions. (See Montgomery, Disadvantages of V-mask)[1]

2.2.2 Construction of Dual-CUSUM Control Chart

Assume the variable of interest X of a particular process happens to follow a standardized normal distribution i.e. $X \sim N(\mu_0 = 0, \sigma = 1)$. The ideal and appropriate CUSUM chart for this process is a one-sided CUSUM chart of the tabular procedure discussed earlier, that uses only equation 1 and plots it against an interval limit H to detect upward shifts in the process mean. The process is expected to always be in an in-control state, except the mean of the process has shifted to say $\mu_1 = \mu_0 + \delta$, that warrants the statistic to be greater than the interval limit, then it's said to be in an out of control state.

Similarly, extending this CUSUM chart to DCUSUM, we have two one-sided classical CUSUM statistics measured and monitored simultaneously as defined below:

$$\left.\begin{aligned} C_{1i}^{+} &= \max\left[0, \left(\hat{\theta}_i - \mu_0\right) - K_1 + C_{1i-1}^{+}\right] \\ C_{2i}^{+} &= \max\left[0, \left(\hat{\theta}_i - \mu_0\right) - K_2 + C_{2i-1}^{+}\right] \end{aligned}\right\}$$

The two statistics are characterized by different parameters, whose selection is discussed below, K_1 & H_1 for C_1^+ and K_2 & H_2 for C_2^+. If $C_1^+ > H_1$ and/or $C_2^+ > H_2$, at any time i, the chart sends a signal and concludes that process is out of control at that time. $C_{10}^+ = C_{20}^+ = 0$, and of course is equivalent to the target mean μ_0. $\hat{\theta}_i$ is a location estimator for the subgroup i of observations of the variable of interest $\{x_{i1}, x_{i2}, \ldots, x_{in}\}$, where n is the sample size for each subgroup. It is important to know that this chart can also be used for variance monitoring.

2.2.3 Selection of the Chart Parameters

In a classical CUSUM charts, the parameters involved are *K* and *H* which are selected for a specified in-control Average Run length (ARL_0), the same way we choose K_1 & H_1 for C_1^+ and K_2 & H_2 for C_2^+ for a specific in-control ARL_0. Meanwhile, Yi et al discussed the conditions and criteria for selecting these parameters without keeping any of the statistic $C_1^+ \; and \; C_2^+$ redundant or ineffective, and by so doing the DCUSUM results to the one sided classical CUSUM.

He argued that, for the DCUSUM chart to be effective the following criteria must be met and not violated in selecting its parameters.

$$\left.\begin{aligned} K_1 &< K_2 \\ H_1 &> H_2 + K_2 - K_1 \\ {}^{K_2}/_{K_1} &\cong {}^{H_1}/_{H_2} \end{aligned}\right\}$$

He intensively justified these conditions logically and algebraically. We let $K_1 < K_2$ such that at any point i, $C_{1i}^{+} > C_{2i}^{+}$ this condition avoids the redundancy of C_{2i}^{+}. Also there exits a point i, where $C_{2i}^{+} > H_2$ which implies $C_{1i}^{+} > C_{2i}^{+} > H_2 > 0$. Recall (from DCUSUM equation) $C_{1i}^{+} - C_{2i}^{+} = K_2 - K_1 + C_{1i-1}^{+} - C_{2i-1}^{+}$, and since $C_{1i}^{+} > C_{2i}^{+}$, $C_{1i-1}^{+} - C_{2i-1}^{+} > 0$, which implies that $C_{1i}^{+} - C_{2i}^{+} \geq K_2 - K_1$ at any point i, and whenever C_{2i}^{+} triggers an out of control signal (i.e. $C_{2i}^{+} > H_2$), implies $C_{1i}^{+} > H_2 + K_2 - K_1$. Now imagine $H_1 \leq H_2 + K_2 - K_1$ this insinuates that whenever C_{2i}^{+} triggers an out-of-control signal, C_{1i}^{+} has already sent a signal as well. Then C_{2i}^{+} is actively working uselessly. Haven justified the first two conditions, there are infinite combinations of K_1, K_2 & H_1, H_2 that meet these conditions. The question is which of them is to be chosen? This question take us back to the asymptotic optimal feature of Lordens' control chart, which simultaneously monitor an infinite number of classical CUSUM statistic defined below:

$$G_{nm} = \sup_{k\epsilon(a/2,b/2)} \{kS_{nm} - (n-m)k^2\}.$$

He (Lordens)[13] proved that this chart has the lowest out-of-control ARL when ARL_0 tends to infinity, meeting the restriction that the multiplication of the parameters of all individual CUSUM charts must be equal i.e. $k_1h_1 = k_2h_2 = k_3h_3 = \cdots k_ih_i = h$. Adopting this restriction for the DCUSUM results to the third condition above.

2.3 DESCRIPTION OF THE PROPOSED ESTIMATORS & THEIR CORRESPONDING DCUSUM CHARTS

An estimator is defined as a function used in calculating an estimate of a parameter based on observed dataset. The estimand (the quantity of interest) could be a location or spread parameter, while the estimator could also be a point or interval type. An estimator, therefore, is said to be robust if it could maintain its performance for a range of dataset drawn from different probability distributions. Hence robust estimators seek to provide resistance to contaminations, outliers and violation of the distribution (under consideration) assumptions which duly affect the method or process.

In this chapter, we shall study the performance of five robust estimators with the DCUSUM chart for the phase II analysis, namely: mean, median, mid-range, Hodges-Lehmann and Tri-mean. These estimators have been studied for the phase I analysis by Schoonhoven et al[17], also studied with classical CUSUM chart for the phase II analysis by Nazir et al [20]

Recall the Dual CUSUM chart defined above, the estimator $\hat{\theta}$ estimating the location parameter θ can be replaced with any of the following estimators. Based on a random sample of size n with observation $\{x_1, x_2, \ldots\ x_n\}$, we define:

1. **Mean $\bar{x}$:** Is the summation of all the observation in a sample divided by the sample size. i.e. $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$. The sample mean is not a robust estimator, though it is a useful and common measure of central tendency, especially for processes that follow a normal distribution, but included in the study as a base of comparison between others.

2. **Median $\tilde{x}$:** Unlike the mean, is a robust estimator of location. It is said to have a breakdown point of 50%, whereas the mean has a breakdown point of 0%. It is the numerical value that separates the higher half of a sample from the lower half. For odd number sample, median is the middle order statistic, and the average of the two middle-order statistic for an even number.
3. **Mid-range:** Is the average of the largest and lowest observation in a sample. It can also be defined as the mid-point of range. Due to its composition, the two extreme values in the sample, makes it very sensitive to outliers. Furthermore, it's also a non-robust estimator as it ignores all other elements of the sample in its computation, having a breakdown point of 0%.
4. **Hodges – Lehmann**: This estimator was independently proposed by Pranab Kumar Sen, Joseph Hodges and Erich Lehmann in 1963[21]. It is also called **Hodges-Lehmann-Sen** estimator. It is defined as the median of the pairwise averages of all the observation in a sample. $HL = median\{(x_i + x_j)/2,\ 1 \le i < j \le n\}$. It is a robust estimator with a breakdown point of 29% and at least 95% of relative efficiency compared to mean with the presence of contaminations and outliers in the probability distribution under consideration. Especially when the probability distribution is normal, the relative efficiency is always above unity.
5. **Tri-Mean:** the tri-mean, also referred to as Turkey tri- mean, is defined as the weighted average of the sample median and its two quantiles. $TM = \frac{Q_1+2Q_2+Q_3}{2}$. It is also equivalent to the average of the median and the mid-hinge $TM = \frac{1}{2}\ (Q_2 + \frac{Q_1+Q_3}{2})$. It is a robust estimator of location that combines the median's emphasis on

center values with the mid-hinge's attention to the extremes. It has a breakdown point of about 25% and 88% efficiency relatively to the sample mean.

Therefore, these proposed estimators of location replaced by $\hat{\theta}$ in the DCUSUM chart

$$\left.\begin{aligned} C_{1i}^{+} &= \max\left[0, \left(\hat{\theta}_i - \mu_0\right) - K_1 + C_{1i-1}^{+}\right] \\ C_{2i}^{+} &= \max\left[0, \left(\hat{\theta}_i - \mu_0\right) - K_2 + C_{2i-1}^{+}\right] \end{aligned}\right\}$$

Constitutes five different DCUSUM charts. Each of them is parameterized by K_1& K_2 and H_1 & H_2. These parameters are carefully selected for a specified ARL.

2.4 PERFORMANCE EVALUATION OF THE PROPOSED DCUSUM CHARTS

In this section, we present the performance of the DCUSUM charts and depict their comparisons. To evaluate the performance of this control charting schemes, we adopted the Monte Carlo Simulation, one of the different methods used in literature for approximation, to obtain the charts' average run lengths (ARL). The in-control ARL always denoted by ARL_0 is the average run length of a process, when there is no shift in the process mean (target). While the out-of control ARL denoted by ARL_1 is the average run length when the process mean has shifted. We also report the standard deviation of the run length (SDRL) to have a better knowledge of the run length distribution.

We developed a code in MATLAB to simulate the ARL and SDRL for the DCUSUM chart described earlier for each of the estimators. The algorithm of the code is such that:

Step 1: generated a 10^5 random numbers each of sample size $n = 5$, from a standard normal distribution.

Step 2: calculated the five estimators in each case, standardized them (haven known the standard deviation of each estimator) and use them as the statistic in the corresponding DCUSUM charts.

Step 3: plotted the charts against their corresponding interval limits H_1 & H_2. At any point the DCUSUM statistics $C_{1i}^{+} > H_1$ and/or $C_{2i}^{+} > H_2$, the process is terminated and the point i is recorded.

Step 4: repeated the procedure 10^5 times, and we found the average and standard deviations of these points which make the ARL and SDRL respectively.

The evaluation and comparisons of these charts will be in different folds. Under the standard normal environments and the contaminated normal environments. The latter is of two types, the location and variance contaminated environments. In this study, each of these are examined with sample size $n = 5$, and an in-control ARL of 370. $i.e. ARL_0 = 370$

2.4.1 Uncontaminated Normal environment

An uncontaminated Normal environment is one which all observation are drawn from $N(\mu_0, \sigma_0^2)$. The ARL's and SDRL's of DCUSUM chart using the different five estimators earlier mentioned under an uncontaminated standard normal distribution is shown in Table 2.1. (pp. 17)

From this Table, we can see that all the five estimators have an in-control (when $\boldsymbol{\delta = 0}$) ARL clustering around 370 which is the target. However, in other cases when $\boldsymbol{\delta \neq 0}$**,** the story changes, though they are all monotonically decreasing as the shift δ increases. It should be noted that the performance of a control chart is evaluated based on how fast it sends an out-of-control signal when the process under monitoring is actually out of control. This can be explained, in the spirit of run length by the lower values of $ARL_1's$ in an out-of-control process. Of course, the sample mean is the best estimator under an uncontaminated normal environment. The Hodges-Lehmann estimator, $HL = median\left\{(x_i + x_j)/2,\ 1 \leq i < j \leq n\right\}$, therefore can be compared to the sample mean, $\overline{\boldsymbol{x}} = \frac{\sum_{i=1}^{n} x_i}{n}$ as it's ARL_1's are almost the same in all cases of shift. Next to these two in performance is the Tri-mean estimator, $TM = \frac{Q_1 + 2Q_2 + Q_3}{2}$ while others are out rightly larger than the sample mean, which is the yardstick of comparison. Furthermore, the SDRL behavior for all the estimators is positively skewed by increment in the shift. Like the ARL, the lower the SDRL, the better the chart as performed in detecting small shifts.

2.4.2 Location – Contaminated Normal Environment (LCNE)

A location Contaminated Normal environment is a model/distribution whose location parameter has been disturbed to a certain percent. Say, A random variable X, is said to be drawn from a LCNE, if it has $(\varphi)100\%$ observations from $N(\mu_0, \sigma_0^2)$ and $(1-\varphi)100\%$ observations from $N(\mu_0 + \ \omega\sigma_0, \sigma_0^2)$ where$-\infty < \omega < \infty$. In this section, we have studied run length distribution of the DCUSUM charts with $\varphi = 0.05$ & 0.01 and $\omega = 4$, with the proposed estimators. In Tables 2.2 and 2.3 (pp. 18 & 19) for $\varphi = 0.05\ and\ 0.01$ respectively, it is obvious that all the five estimators are not adequate for the targeted ARL_0

as they are drastically lower than $ARL_0 = 370$. Even with the poor performances, we can still compare the estimators based on their resistance and robustness to the disturbances.

In both cases, when , median turns to have the highest ARL_0, followed by tri-mean, Hodges-Lehmann. The remaining estimators are therefore not comparable with the first three due to their low performances. It should be noted that, the substantial drop in the ARL_0 values of these estimators is because they are measures of location and the disturbance is also on the location parameter. This can be justified by the increase in the values of ARL_0's for all the estimators, as the level of disturbance φ decreases. In Table 2.3, when $\varphi = 0.01$, the ARL_0's values of the estimators are although not comparable with that of the uncontaminated normal environment, but substantially greater compared to when $\varphi = 0.05$. This poor performance in the ARL_0, thwarted the aim of analyzing the $ARL_1's$ results and comparing the estimators based on these values.

2.4.3 Variance – Contaminated Normal Environment (VCNE)

The VCNE like the LCNE, is a distribution whose spread parameter has been disturbed. A random variable say X, is said to be drawn from a VCNE, if it has $(\varphi)100\%$ observations from $N(\mu_0, \sigma_0^2)$ and $(1-\varphi)100\%$ observations from $N(\mu_0, \tau\sigma_0^2)$ where$-\infty < \tau < \infty$. Similarly, we studied the performances of the proposed estimators with their corresponding charts under this environment. We examined when $\varphi = 0.05$ & 0.01 and $\tau = 9$. Table 2.4 and 2.5 show the ARL and SDRL values for both in-control and out-of-control cases.

Generally, all the estimators under the variance-contaminated environment performed better compared to the location-contaminated environment because the estimators are estimating location parameter. Nonetheless, under $ARL_{0,}$ the sample median out performs

other estimators, as it has the closest value to the target, especially when the percent of contamination is low. Next to the median, is the $TM\ and\ HL$, which have comparable values to the median.

Table 2.0: ARL and SDRL Values for the DCUSUM Based on Different estimators Under an Uncontaminated Standard Normal Environment with $k_1 = 0.12,\ k_2 = 0.50$ and different h_1 & h_2 for a specified $ARL_0 = 370$

		δ											
$\widehat{\theta}$	*Value*	**0**	**0.25**	**0.33**	**0.4**	**0.48**	**0.55**	**0.63**	**0.7**	**0.78**	**0.85**	**0.93**	**1**
$\bar{x}$	***ARL***	372.81	22.241	14.044	10.268	7.7818	6.3805	5.3	4.624	4.037	3.6363	3.2807	3.0216
	SDRL	366.51	15.705	9.2641	6.1124	4.1266	3.0576	2.3113	1.8687	1.5181	1.2909	1.1052	0.9648
$\tilde{x}$	***ARL***	369.54	29.991	19.244	13.88	10.289	8.2921	6.7454	5.8084	5.0083	4.4824	3.9915	3.6604
	SDRL	362.78	21.695	13.575	9.2983	6.2849	4.6148	3.4155	2.727	2.1957	1.8474	1.535	1.3493
MR	***ARL***	374.02	27.524	17.573	12.776	9.5206	7.7283	6.3822	5.5279	4.7833	4.2927	3.8563	3.536
	SDRL	366.21	19.738	12.094	8.1951	5.4603	4.0428	3.0208	2.4286	1.9538	1.6499	1.4025	1.2262
TM	***ARL***	371.81	24.526	15.230	11.126	8.3581	6.8227	5.6074	4.8581	4.2616	3.7989	3.4041	3.1483
	SDRL	367.01	17.936	10.417	6.9025	4.7911	3.4796	2.6163	2.0483	1.6784	1.4322	1.1836	1.0484
HL	***ARL***	369.67	23.070	14.549	10.757	8.0994	5.5992	5.5038	4.7994	4.1852	3.7563	3.4069	3.1277
	SDRL	350.16	16.569	9.6403	6.4507	4.34	3.1832	2.4452	2.0088	1.5884	1.3611	1.168	1.0337

Table 2.1: ARL and SDRL Values for the DCUSUM Based on Different estimators Under 5% LCNE with $k_1 = 0.12,\ k_2 = 0.50$ and different h_1 & h_2 for a specified $ARL_0 = 370$

		δ											
$\widehat{\theta}$	*Value*	**0**	**0.25**	**0.33**	**0.4**	**0.48**	**0.55**	**0.63**	**0.7**	**0.78**	**0.85**	**0.93**	**1**
$\overline{x}$	***ARL***	20.836	8.2625	6.7654	5.7188	4.9192	4.3587	3.8704	3.4875	3.1819	2.9516	2.7239	2.578
	SDRL	17.658	5.4975	4.0873	3.2021	2.5891	2.0832	1.7655	1.4917	1.2847	1.1401	1.0197	0.9326
$\widetilde{x}$	***ARL***	81.452	17.902	12.879	10.128	8.001	6.7019	5.7429	5.0299	4.4598	4.0257	3.6365	3.3577
	SDRL	75.114	13.386	9.0078	6.5406	4.6048	3.5683	2.8778	2.3273	1.948	1.6707	1.4117	1.2545
MR	***ARL***	11.107	6.7267	5.8873	5.2579	4.6611	4.2697	3.8554	3.5823	3.3013	3.091	2.8623	2.7156
	SDRL	8.9646	4.7053	3.9582	3.3589	2.7762	2.4743	2.0775	1.853	1.5991	1.4319	1.2684	1.1794
TM	***ARL***	57.714	13.931	10.373	8.2829	6.8083	5.8289	4.9789	4.4346	3.9522	3.6201	3.2955	3.0402
	SDRL	49.227	9.7893	6.6546	4.8805	3.5868	2.9045	2.2589	1.8616	1.5575	1.3756	1.1914	1.0524
HL	***ARL***	33.787	10.881	8.3406	6.9275	5.7818	5.0125	4.3872	3.9213	3.534	3.2623	2.9768	2.7723
	SDRL	30.081	7.5711	5.2414	4.0607	3.0798	2.4593	2.0117	1.6789	1.4346	1.2381	1.0808	0.9614

Table 2.2: ARL and SDRL Values for the DCUSUM Based on Different estimators Under 1% LCNE with $k_1 = 0.12,\ k_2 = 0.50$ and different h_1 & h_2 for a specified $ARL_0 = 370$

		δ											
$\widehat{\theta}$	*Value*	**0**	**0.25**	**0.33**	**0.4**	**0.48**	**0.55**	**0.63**	**0.7**	**0.78**	**0.85**	**0.93**	**1**
$\bar{x}$	***ARL***	116.74	16.787	11.542	8.8323	6.9545	5.8693	4.9114	4.3147	3.8281	3.5044	3.1565	2.9094
	SDRL	109.66	12.129	7.3911	5.1794	3.7375	2.829	2.1972	1.7678	1.4837	1.2777	1.0894	0.9571
$\tilde{x}$	***ARL***	263.66	27.087	17.390	12.985	9.7664	7.9887	6.5273	5.5808	4.8782	4.3727	3.9358	3.6108
	SDRL	258.89	19.810	12.048	8.588	5.8981	4.437	3.3	2.5853	2.1187	1.7906	1.5413	1.3164
MR	***ARL***	56.172	16.737	12.168	9.7627	7.7859	6.6184	5.6385	4.9826	4.3735	3.973	3.6185	3.3326
	SDRL	52.159	12.912	8.5849	6.2771	4.599	3.598	2.8033	2.3053	1.9036	1.6402	1.4036	1.241
TM	***ARL***	201.46	20.383	13.591	10.186	7.7621	6.3777	5.3369	4.6664	4.1099	3.7145	3.3272	3.0534
	SDRL	195.27	14.903	9.0882	6.3149	4.3218	3.1781	2.4271	2.004	1.6075	1.4013	1.1787	1.0282
HL	***ARL***	175.89	19.555	12.865	9.69	7.5162	6.2791	5.1993	4.6452	4.0212	3.631	3.2938	2.0563
	SDRL	166.17	13.995	8.6236	5.8067	4.0048	3.0978	2.3189	1.9351	1.5574	1.3328	1.1365	1.0107

Table 2.3: ARL and SDRL Values for the DCUSUM Based on Different estimators Under 5% DCNE with $k_1 = 0.12,\ k_2 = 0.50$ and different h_1 & h_2 for a specified $ARL_0 = 370$

		δ											
$\widehat{\theta}$	*Value*	**0**	**0.25**	**0.33**	**0.4**	**0.48**	**0.55**	**0.63**	**0.7**	**0.78**	**0.85**	**0.93**	**1**
$\bar{x}$	***ARL***	122.51	18.960	12.877	9.9174	7.5762	6.3055	5.2822	4.6283	4.0624	3.6617	3.3213	3.0507
	SDRL	118.61	14.683	9.1043	6.4484	4.3576	3.3546	2.563	2.1188	1.7049	1.472	1.251	1.1027
$\tilde{x}$	***ARL***	260.12	28.105	18.474	13.602	10.201	8.2345	6.6891	5.8258	5.0413	4.4905	3.988	3.6544
	SDRL	254.91	21.198	13.313	9.3115	6.4241	4.7119	3.4778	2.8919	2.2705	1.9075	1.6058	1.3781
MR	***ARL***	57.779	18.937	14.148	11.232	8.8488	7.4493	6.2269	5.4707	4.7904	4.338	3.8645	3.5508
	SDRL	56.259	15.461	10.815	8.0139	5.8504	4.5164	3.456	2.8943	2.3869	2.014	1.7301	1.5013
TM	***ARL***	217.14	22.401	14.619	10.913	8.2954	6.7609	5.5626	4.8322	4.2455	3.4376	3.1605	3.1605
	SDRL	211.13	16.769	10.314	7.1073	4.7864	3.5556	2.6769	2.1733	1.7625	1.2983	1.1273	1.1273
HL	***ARL***	194.28	21.347	14.076	10.489	8.0206	6.5523	5.4602	4.7768	4.2108	3.7457	3.3928	3.1242
	SDRL	188.34	16.116	9.8233	6.7415	4.611	3.3259	2.5687	2.0733	1.7274	1.4388	1.2204	1.0933

Table 2.4: ARL and SDRL Values for the DCUSUM Based on Different estimators Under 1% DCNE with $k_1 = 0.12,\ k_2 = 0.50$ and different h_1 & h_2 for a specified $ARL_0 = 370$

		δ											
$\widehat{\theta}$	*Value*	**0**	**0.25**	**0.33**	**0.4**	**0.48**	**0.55**	**0.63**	**0.7**	**0.78**	**0.85**	**0.93**	**1**
$\overline{x}$	***ARL***	268.11	21.331	13.737	10.261	7.7516	6.3504	5.2843	4.643	4.045	3.6535	3.2812	3.0351
	SDRL	261.94	15.297	9.2306	6.2725	4.1764	3.1477	2.3604	1.9299	1.5471	1.3447	1.13	1.0176
$\widetilde{x}$	***ARL***	343.24	30.008	19.071	13.761	10.357	8.298	6.7334	5.8343	5.0335	4.4728	4.0026	3.6431
	SDRL	342.13	21.608	13.509	9.1627	6.2999	4.6993	3.4473	2.7794	2.2118	1.8542	1.539	1.343
MR	***ARL***	179.40	25.404	16.994	12.553	9.4718	7.6576	6.2897	5.5034	4.7557	4.2986	3.8729	3.5417
	SDRL	18.856	18.856	12.075	8.2925	5.6628	4.1621	3.141	2.5046	2.04	1.7575	1.4993	1.2814
TM	***ARL***	328.50	23.940	15.244	11.297	8.332	6.8284	5.6202	4.8689	4.2296	3.8148	3.4201	3.1379
	SDRL	320.35	17.468	10.497	7.2553	4.6147	3.4694	3.4694	2.123	1.6968	1.4335	1.2264	1.0479
HL	***ARL***	321.18	22.826	14.554	10.747	8.0993	6.6197	5.4654	4.7985	4.1815	3.7587	3.39	3.1249
	SDRL	318.55	16.394	9.8057	6.583	4.3851	3.314	2.4571	1.6517	1.6517	1.3668	1.1965	1.0331

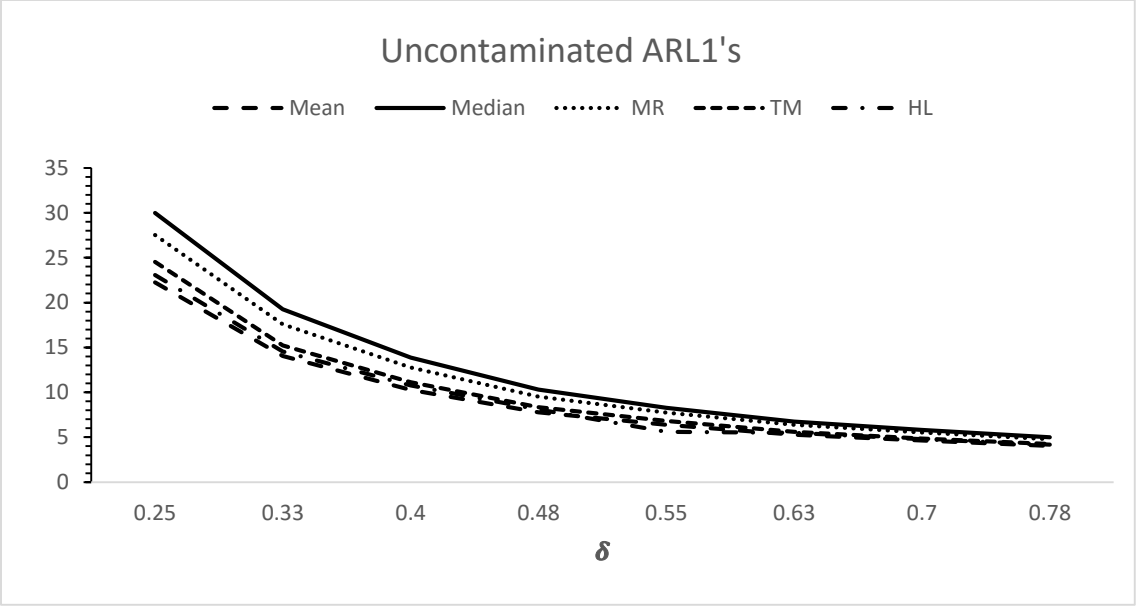

Figure 2.0: The ARL_1's curve of the uncontaminated environment

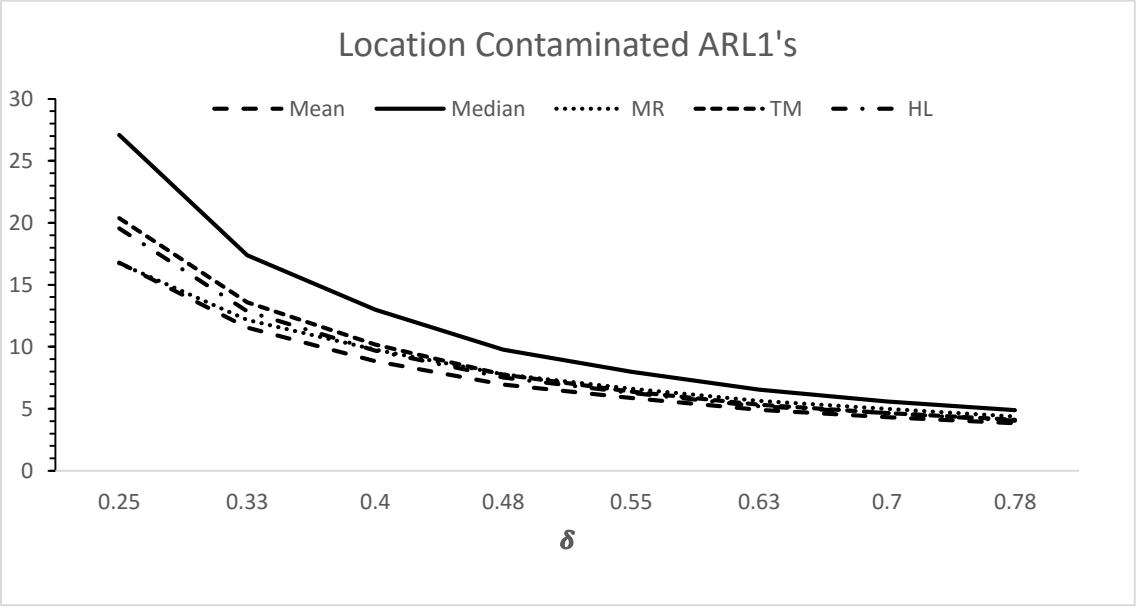

Figure 2.1: The ARL_1's curve of the location-contaminated environment

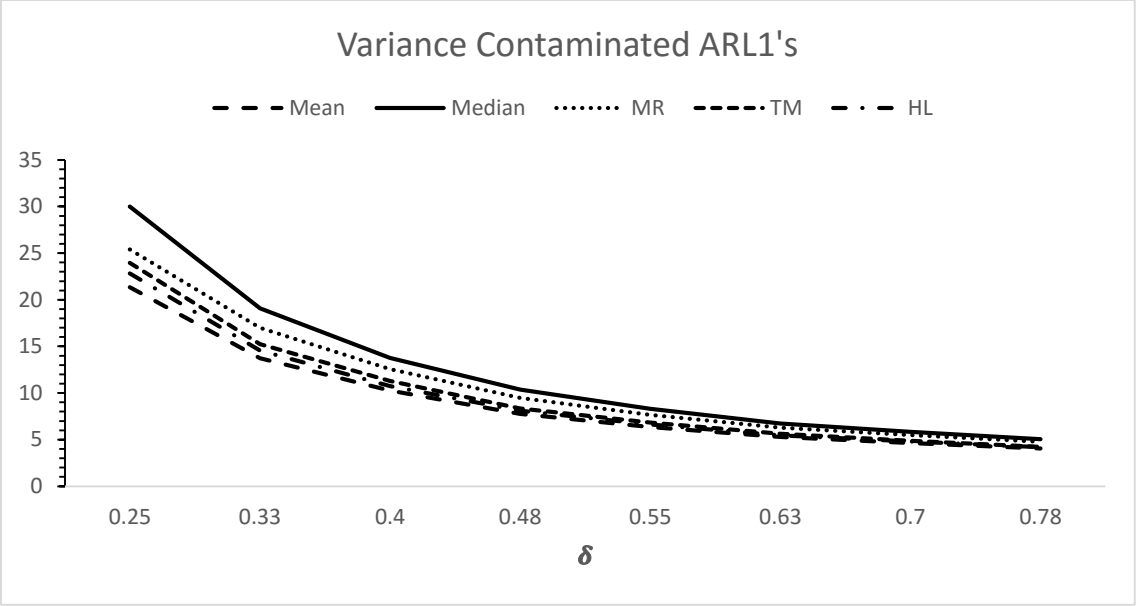

Figure 2.2: The ARL_1's curve of the variance-contaminated environment

Figure 2.3: The ARL_0 values of the uncontaminated environment

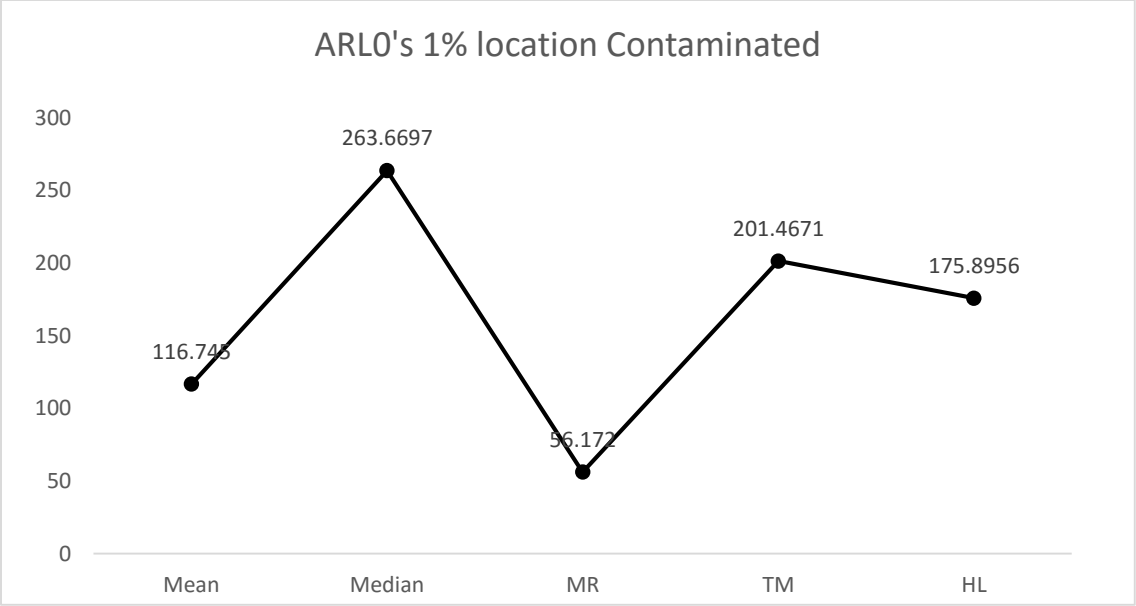

Figure 2.4: The ARL_0 values of the variance-contaminated environment

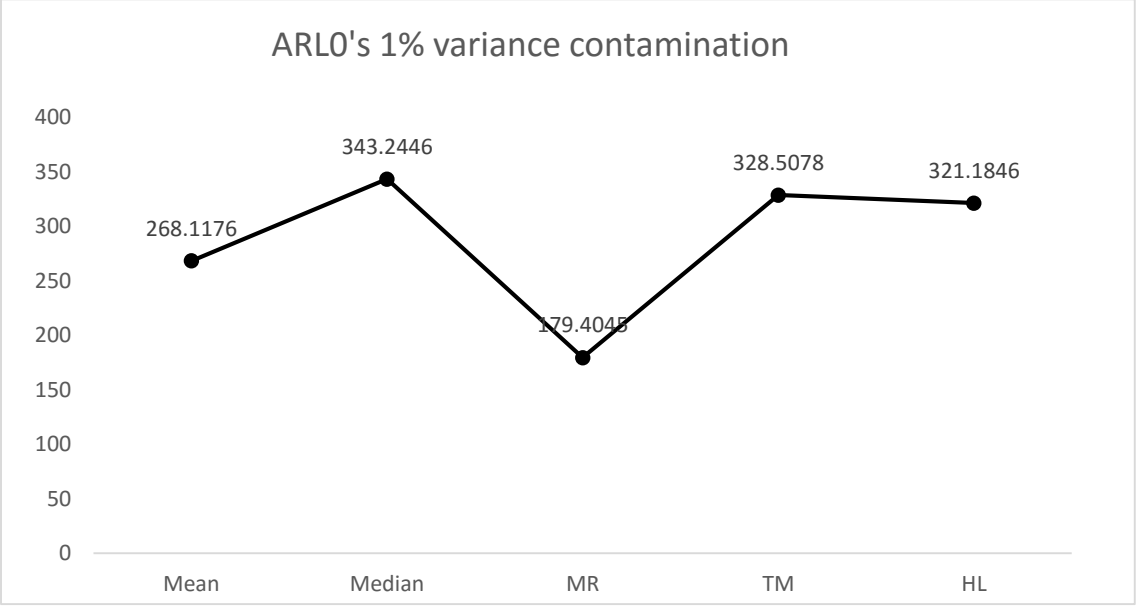

Figure 2.5: The ARL_0 values of the variance-contaminated environment

2.5 APPLICATION OF THE SCHEMES ON A REAL LIFE DATASET

Claims without proofs are mere expressions they say. We have in this section, substantiate the efficiency of these charts by a practical demonstration with a real life data set. The data set which consists of 30 sample points with the sample size n = 5, is the waiting time in minutes of patients for a colonoscopy procedure in a regional health center[22]. The original data set without contamination is presented in Figure 2.6 below, while the subsequent Figures 2.8-2.12 show the DCUSUM control chart of respective charts. Though, in real-life we could not predict where the disturbances could lie, it could be on the location, variance or both location and variance parameters. For all the estimators, their respective charts behave quite well under the uncontaminated normal environment as we can see all the points of the charts falling within the control limits.

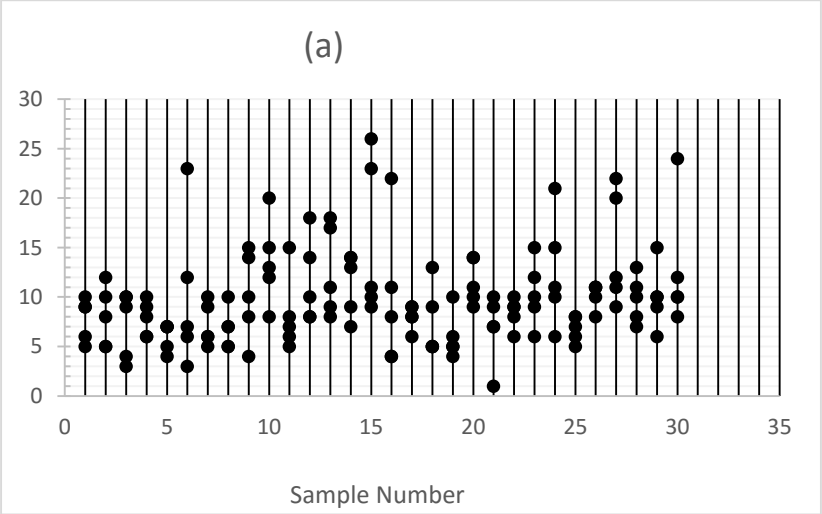

Figure 2.6: The scatter plot of the uncontaminated dataset

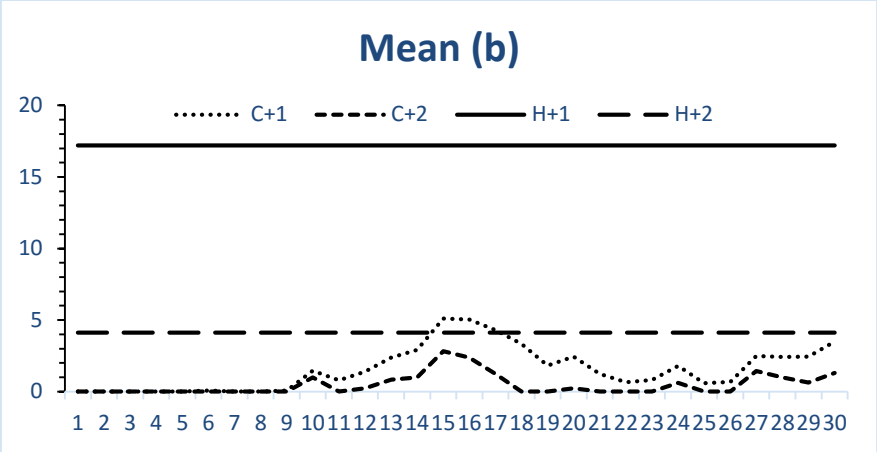

Figure 2.7: Control Chart of mean DCUSUM statistics for uncontaminated dataset

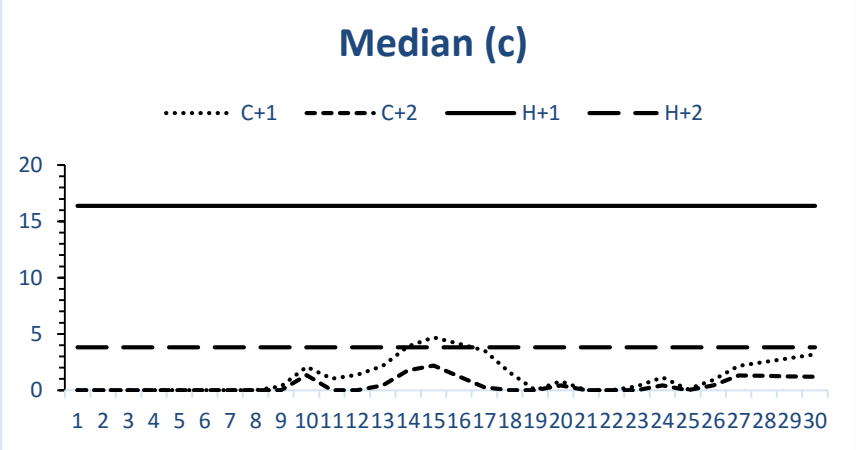

Figure 2.8: Control Chart of median DCUSUM statistics for uncontaminated dataset

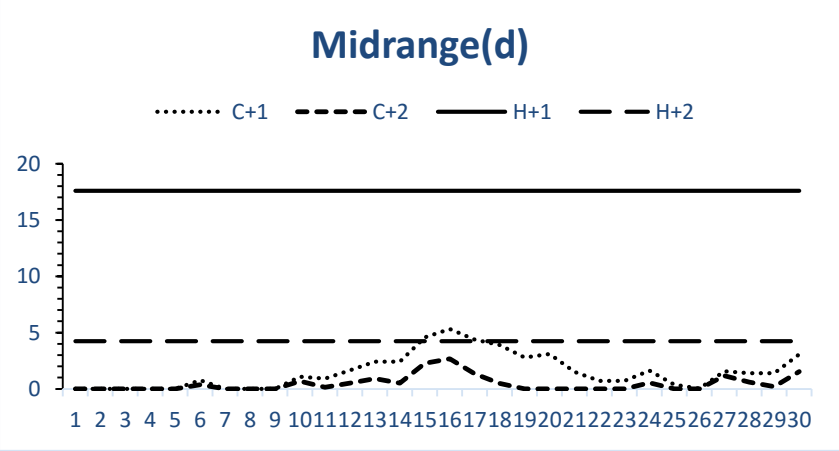

Figure 2.9: Control Chart of midrange DCUSUM statistics for uncontaminated dataset

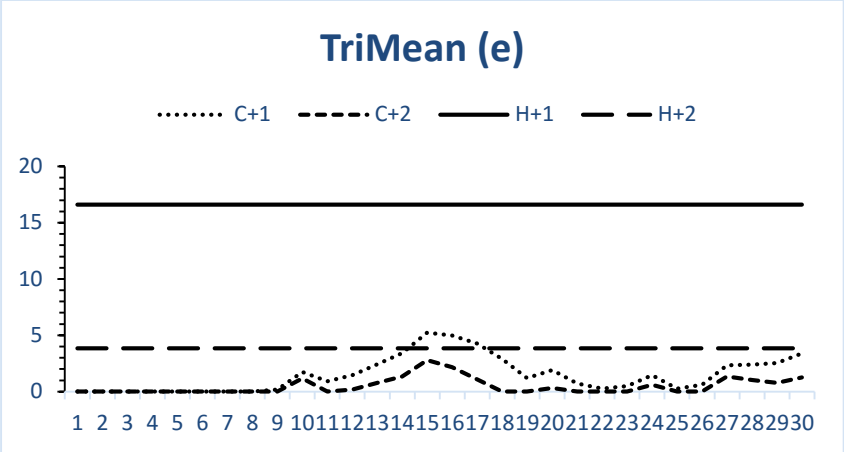

Figure 2.10: Control Chart of trimean DCUSUM statistics for uncontaminated dataset

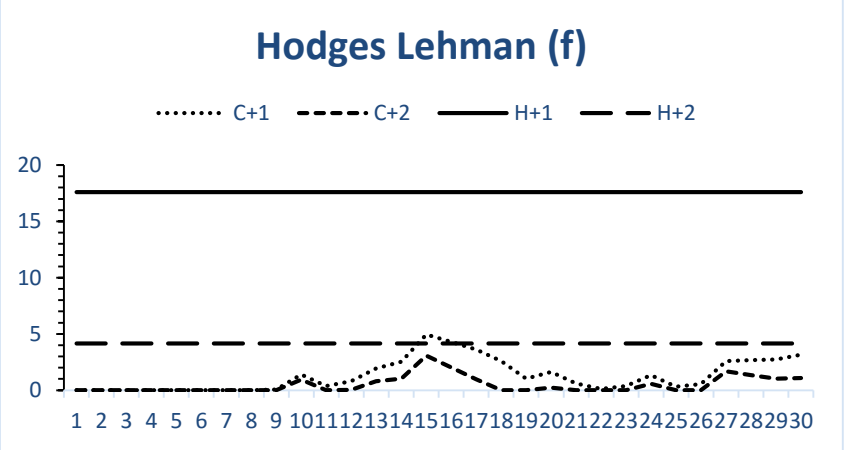

Figure 2.11: Control Chart of HL DCUSUM statistics for uncontaminated

For the location-contaminated environments, comes the survival of the fittest rule, as we can see from Figures 2.14 and 2.16 that both mean and midrange could not withstand the location contamination applied to the original data set. At sample point 14, the midrange chart went out of its limit until point 17, which makes it the least robust of all the estimator. Next to it in bad performance, is the mean, as it also got to an out –of control stage at 15 and maintains its limit back immediately at point 16.

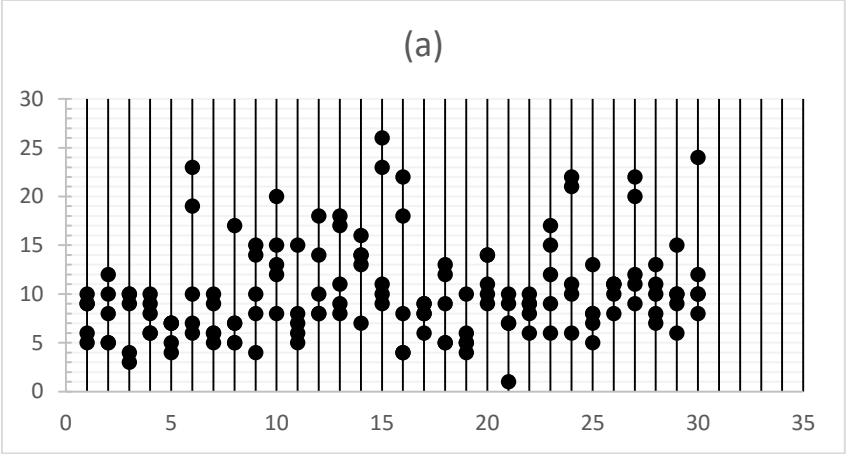

Figure 2.12: The scatter plot of the LCNE dataset

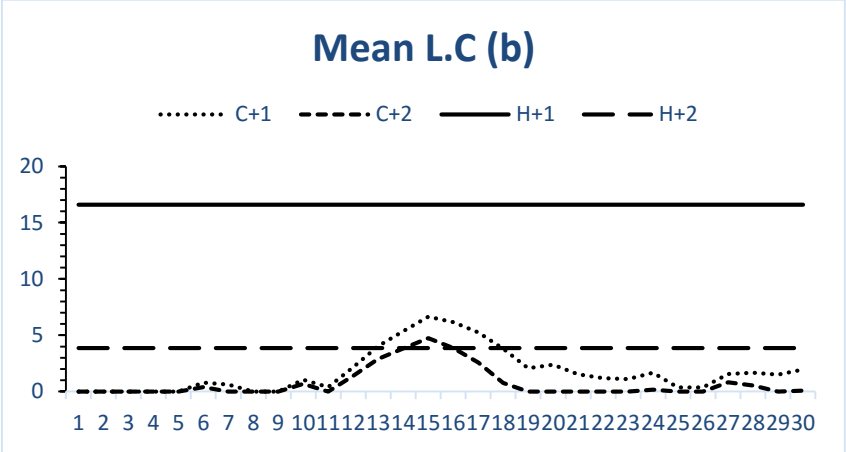

Figure 2.13: Control Chart of mean DCUSUM statistics for LCNE dataset

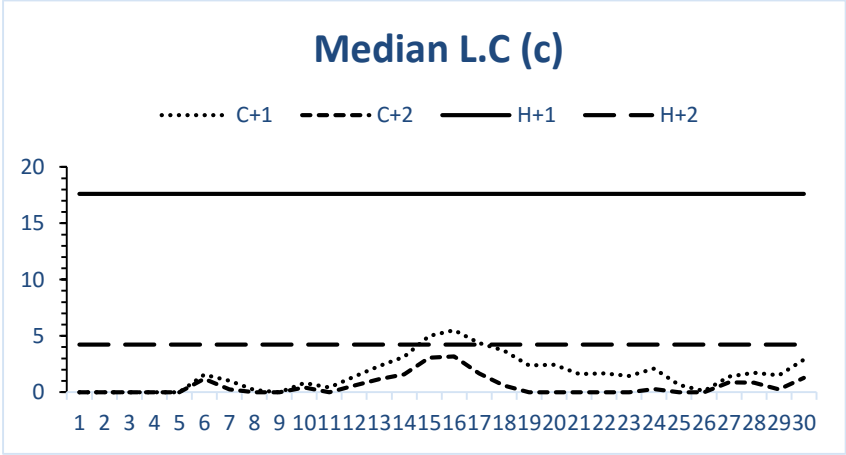

Figure 2.14: Control Chart of median DCUSUM statistics for LCNE dataset

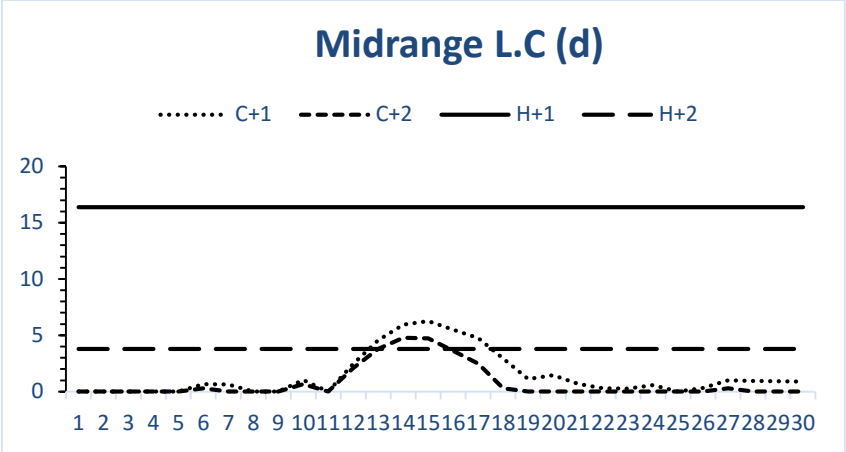

Figure 2.15: Control Chart of midrange DCUSUM statistics for LCNE dataset

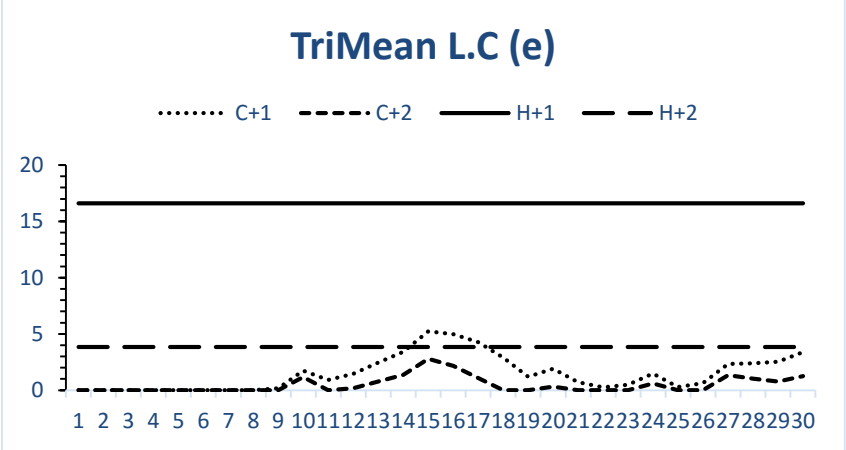

Figure 2.16: Control Chart of trimean DCUSUM statistics for LCNE dataset

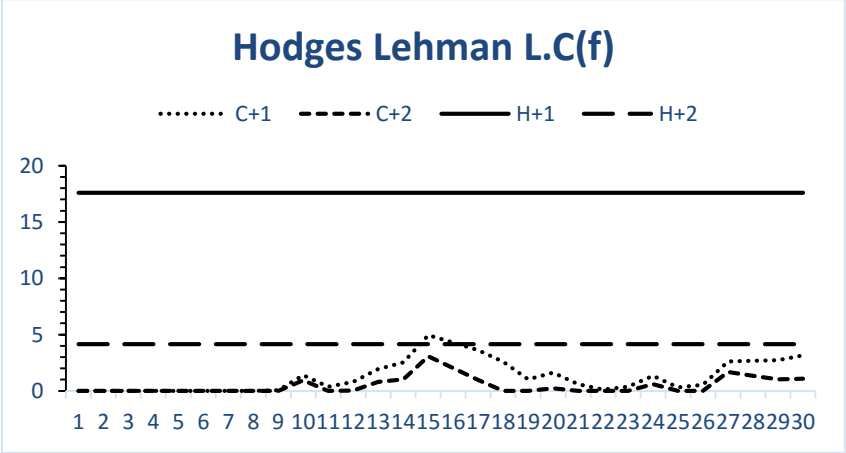

Figure 2.17: Control Chart of HL DCUSUM statistics for LCNE dataset

The evaluation of these schemes on the real-life data set is based on how much each estimator could withstand the respective location and variance contaminations. The estimator whose chart gets an out-of-control state (by getting beyond its limit) will be declared as less-robust estimator. All the estimators in the normal environment without contamination did not send signals which implies all the estimators are robust. Out of the five estimators, under the location contamination, the median chart proves to be the best, as we can see from the charts, followed by trimean and Hodges-Lehmann estimators hierarchically. The real-life example as just buttressed the simulation result produced in that regard.

The same scenario occurs under the variance contaminated environment sample, the median retains its position as the best of all, followed by trimean and Hodges-Lehmann chart respectively. However, the performance of the median and mean charts under the variance contamination seems to be better than that of location contamination, but relatively, the two charts are still in an out-of-control state at some points, as it's obvious in their respective charts.

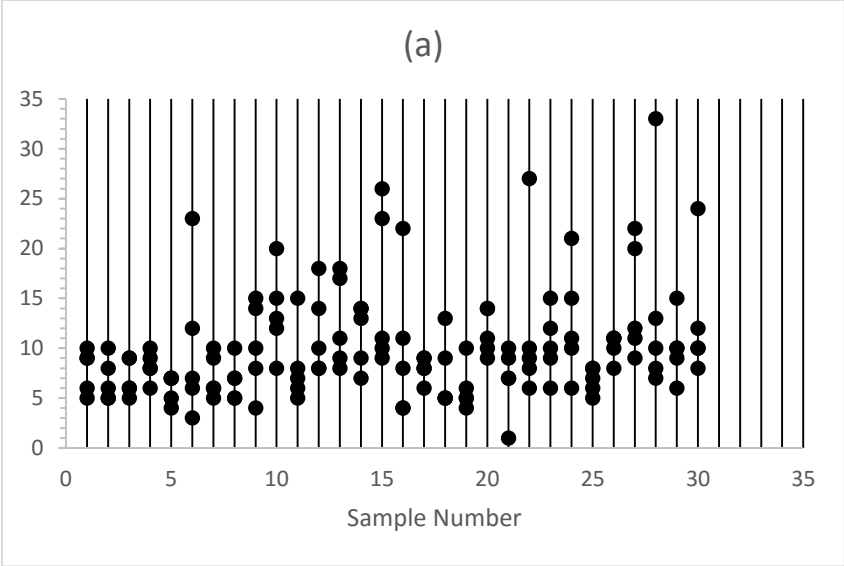

Figure 2.18: The scatter plot of the variance-contaminated dataset

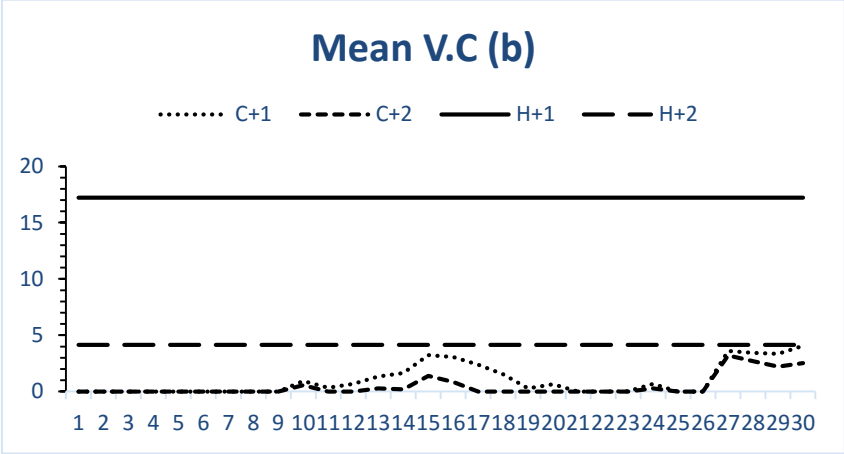

Figure 2.19: Control Chart of mean DCUSUM statistics for VCNE dataset

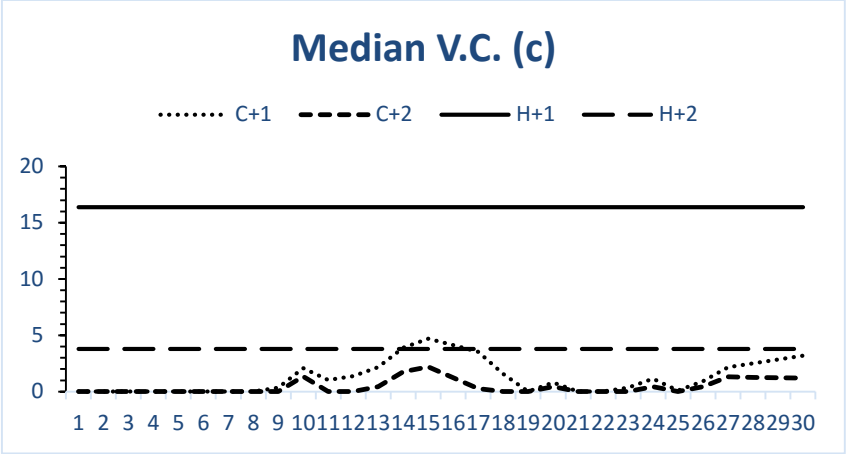

Figure 2.20: Control Chart of median DCUSUM statistics for VCNE dataset

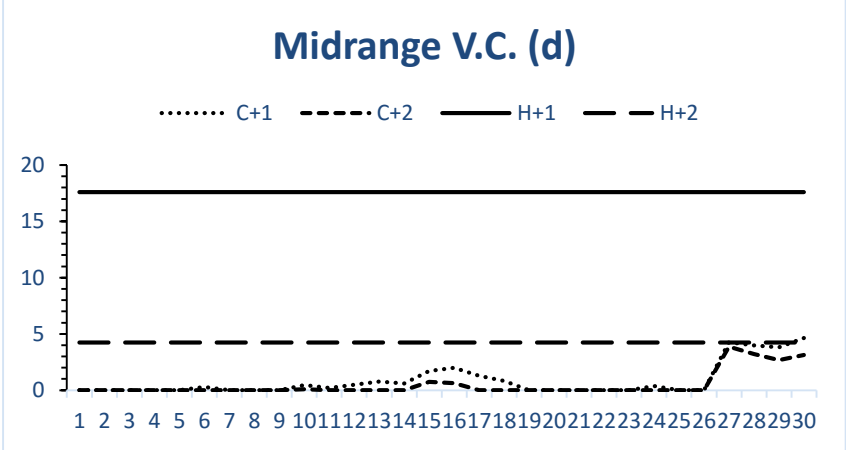

Figure 2.21: Control Chart of midrange DCUSUM statistics for VCNE dataset

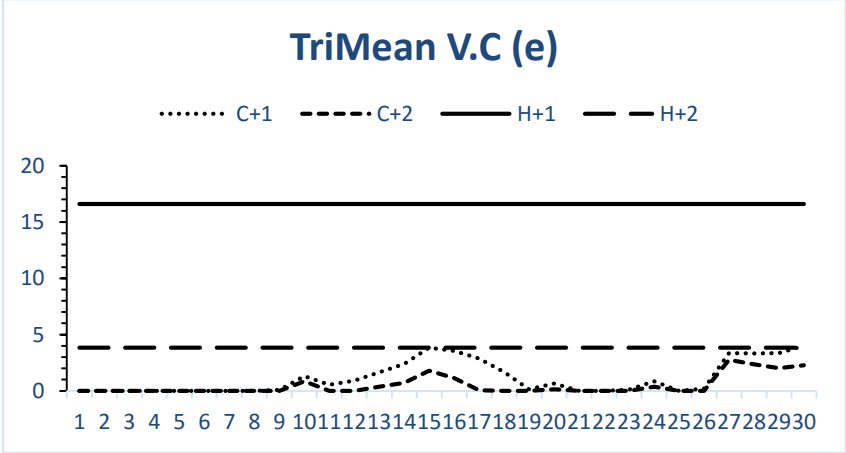

Figure 2.22: Control Chart of trimean DCUSUM statistics for VCNE dataset

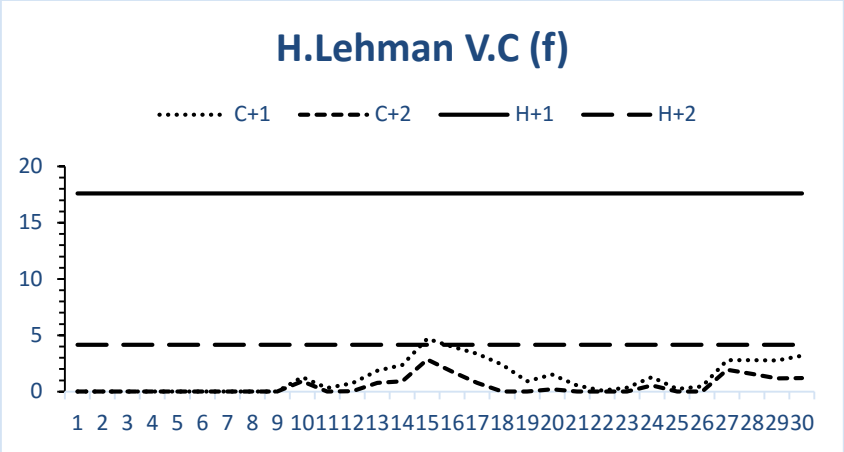

Figure 2.23: Control Chart of HL DCUSUM statistics for VCNE dataset

2.6 SUMMARY AND CONCLUSION

Control charts are the widely used tool of SPC in for monitoring manufacturing and service processes. A memory type of this chart is the classical CUSUM chart that is efficient in detecting small and moderate shifts in any process. The one-sided DCUSUM chart has been proved to have better performance compared to the classical CUSUM in terms of the Average Run Length (ARL), while the former was designed to detect a range of varying shifts in a process, the latter was designed to monitor a specified shift in a process. In this study we establish that robust DCUSUM charts are preferred to the ordinary DCUSUM charts as robust classical CUSUM were preferred to classical CUSUM. Adopting some robust statistics in the presence of violations of normality assumptions under normal environments such as contaminations and outliers, makes the corresponding charts maintain their efficient and robustness. Amongst the five location estimators used in designing the DCUSUM charts, with the results shown above, the mean estimator happens to be the most efficient of all under non- contaminated normal environments, while the TM and HL estimators are good comparisons to the mean. In the presence of location

contamination, the performance of all the estimators are quite unsatisfactory, which is expected, because the estimators are estimating location parameters. Despite this worst performance, the TM and HL estimators' charts outperform others, based on ARL_0 and ARL_1 examination. Though the ARL_0 of the median seems to be the best, but the median chart performance is slow compared to TM and HL performances in detecting small shifts in the process which is of paramount. The performance of these charts under dispersion contamination environments depicts the efficiency and robustness of the TM and HL estimators better.

In general, this study substantially proves that the DCUSUM chart is preferred to the ordinary CUSUM and TM estimator is the best amongst the five estimators, under all circumstances.

CHAPTER 3
ON DESIGNING A ROBUST DOUBLE-EWMA CONTROL CHART FOR PROCESS MONITORING

3.1 INTRODUCTION

Since its advent by Robert[4], the exponentially weighted moving average EWMA chart has gained maximum attention of researchers both in academia and industry. EWMA chart as a type of quality control chart in statistical process control, and its counterpart Cumulative Sum chart fall under memory charts unlike the Shewhart chart that belongs to Memoryless type. The memory type does not ignore the past information like the memory-less, rather it combines both immediate and past information about the process in constructing their respective charts. This feature of the memory type chart (EWMA for example) makes it very sensitive in detecting small and moderate drifts in any process.

EWMA chart like other control charts is of two phases; in phase I, we primarily construct the chart, design the control limits and make sure the process is in a state of statistical control over a period of time, while in the phase II, we use the designed charts therein phase I to monitor the process in comparison with the in-control process we had in phase I. Generally, most control charts are designed without paying much attention to the phase I, with the mindset that all processes obey the assumption of normality and by that their respective parameters are known. Meanwhile, these parameters should be estimated whenever they are unknown and normality assumption is being violated.

In line of increasing the sensitivity of EWMA chart, numerous modifications have been made and published in the literature, to mention a few; the omnibus EWMA method by Domangue and Patch [9], adaptive EWMA chart by Capizi and Masarotto [10], also Borror

et al [11], extensively investigated the performance of EWMA chart under non-normal distributions. In line of extending the classical EWMA to double EWMA, Shamma and Shamma [23] and Zhang [24] have independently designed a double EWMA (DEWMA) control charts for mean, Mahmoud and Woodall [25] also evaluated the signal resistance measure of DEWMA chart, emphasizing its superiority over the classical EWMA chart based on zero-state and worse case average run length (ARL) measures. Similarly, Alkahtani [26] wrote on robustness of DEWMA against EWMA control charts under non-normal processes.

In this chapter, we aimed at working on the efficiency of the DEWMA control charts by proposing some robust estimators of location parameter in constructing the charts, also analyzing the effect of parameter estimation on the performance of these proposed charts at the phase I stage where the parameters are unknown. Since the superiority of DEWMA against EWMA has been established, we shall compare the performance of these estimators within themselves and see which of the estimators retains its robustness in the presence of contaminations and disturbances. All these are done with a synthetic data through Monte-Carlo simulations and then buttressed with its application in a real life data set.

The next section of this chapter describes the classical/standard EWMA chart and explains the DEWMA chart: its designation and types. While section 3 is a detailed description of the proposed DEWMA charts section 4 depicts the performance evaluation of the proposed charts, and the effects of parameter estimation in phase I of these charts with some simulation results follows in section 5. We present an application of the charts with a real-life data set in section 6, finally comes the summary and conclusion in section 7 of the chapter.

3.2 EWMA & DEWMA control charts

The memory-type control chart EWMA for process mean which was introduced by Robert [4] is defined as follows: Suppose a random variable X_i, where $i = 1,2, \dots, n$ is a sequence of random variables with mean μ_0 and variance σ^2, the EWMA statistic is given as:

$$Z_i = \lambda X_i + (1-\lambda) Z_{i-1}$$

where $0 < \lambda \leq 1$ is a smoothing constant and $Z_0 = \mu_0$. The smaller the value of λ the more sensitive the chart becomes, if $\lambda = 1$, the chart becomes a Shewhart chart, which is memory-less and suitable for large shifts. When dealing with subgroup data (i.e. $m > 1$), we replace X_i by any estimator θ_i and every other thing follows suit. Therefore, the EWMA control chart statistic Z_i is plotted against the sample (or subgroup) number i with the center line, upper and lower limits given below:

$$\left.\begin{aligned} CL &= \mu_0 \\ UCL &= \mu_0 + L\sigma\sqrt{\frac{\lambda}{(2-\lambda)}\left[1-(1-\lambda)^{2i}\right]} \\ LCL &= \mu_0 - L\sigma\sqrt{\frac{\lambda}{(2-\lambda)}\left[1-(1-\lambda)^{2i}\right]} \end{aligned}\right\}$$

where L is the distance between the control limits and the center line (CL) measured in σ units. It should be noted here that the term $\left[1-(1-\lambda)^{2i}\right]$ goes to unity as i increases. Then the resulting UCL and LCL are:

$$UCL = \mu_0 + L\sigma\sqrt{\left(\frac{\lambda}{(2-\lambda)}\right)} \text{ and } LCL = \mu_0 - L\sigma\sqrt{\left(\frac{\lambda}{(2-\lambda)}\right)}.$$

On the other hand, the double exponentially weighted moving average (DEWMA) chart is an extension of the EWMA to increase its sensitivity of detecting smaller shifts. This idea was pioneered by Brown [27]to forecast future time series observation. The i^{th} DEWMA chart statistic is given as:

$$\left.\begin{aligned} Z_i &= \lambda_1\hat{\theta}_i + (1-\lambda_1)Z_{i-1} \\ W_i &= \lambda_2 Z_i + (1-\lambda_2)W_{i-1}, i = 1,2,\dots, \end{aligned}\right\}$$

where $Z_0 = W_0 = \mu_0$. The choice of the smoothing parameters (λ_1&λ_2) constitutes the types of DEWMA, one could make $\lambda_1 = \lambda_2$ or $\lambda_1 \neq \lambda_2$. Whichever way, Zhang & Chen [27]concluded that the DEWMA chart with equal smoothing factors does not out performs the other with different smoothing factors and vice-versa. They showed the exact variance of the chart statistic for both cases to be:

$$\sigma^2_{Z_i,exact} = \begin{cases} \dfrac{\lambda_1^2\lambda_2^2}{(\lambda_1-\lambda_2)^2}\left[\dfrac{(\lambda_2')^2\left(1-\lambda_2'^{2i}\right)}{1-\lambda_2'^2} + \dfrac{(\lambda_1')^2\left(1-\lambda_1'^{2i}\right)}{1-\lambda_1'^2} - 2\dfrac{\lambda_1'\lambda_2'\left(1-(\lambda_1'\lambda_2')^i\right)}{1-\lambda_1'\lambda_2'}\right]\sigma_0^2 \\ \qquad if\ \lambda_1 \neq \lambda_2 \\ \lambda_1^4\dfrac{1+(\lambda_1')^2-(i^2+2i+1)(\lambda_1')^{2i}+(2i^2+2i+1)(\lambda_1')^{2i+2}-\ i^2(\ (\lambda_1')^{2i+4}}{(1-(\lambda_1')^2)^3}\sigma_0^2 \\ \qquad if\ \lambda_1 = \lambda_2 \end{cases}$$

where $\lambda_1' = (1-\lambda_1)$ and $\lambda_2' = (1-\lambda_2)$. While the asymptotic variance of the DEWMA statistic was shown by Mahmoud & Woodall [25] to be:

$$\sigma^2_{Z_i,asymptotic} = \begin{cases} \dfrac{\lambda_1^2\lambda_2^2}{(\lambda_1-\lambda_2)^2}\left[\dfrac{(\lambda_2')^2}{1-\lambda_2'^2} + \dfrac{(\lambda_1')^2}{1-\lambda_1'^2} - 2\dfrac{\lambda_1'\lambda_2'}{1-\lambda_1'\lambda_2'}\right]\sigma_0^2 & if\ \lambda_1 \neq \lambda_2 \\ \lambda_1^4\dfrac{1+(\lambda_1')^2}{(1-(\lambda_1')^2)^3}\sigma_0^2 & if\ \lambda_1 = \lambda_2 \end{cases}$$

Therefore, the control limits for the DEWMA charts are with their respective cases:

Exact	Asymptotic
$CL = \mu_0$	$CL = \mu_0$
$UCL = \mu_0 + L\sqrt{\sigma^2_{Z_i,exact}}$	$UCL = \mu_0 + L\sqrt{\sigma^2_{Z_i,asymptotic}}$
$LCL = \mu_0 - L\sqrt{\sigma^2_{Z_i,exact}}$	$LCL = \mu_0 - L\sqrt{\sigma^2_{Z_i,asymptotic}}$

The DEWMA statistic like the EWMA chart, is plotted against its control limits, and at any point i where the Z_i gets beyond either of the two bounds, the process is said to be in an out-of-control state, otherwise, it is in-control.

3.3 DESCRIPTION OF THE PROPOSED DEWMA CHART

Without any loss of generality and misconception of terms, we shall make some alteration in the afore-described DEWMA charts, by replacing $X_i$with some estimator because we shall be dealing with subgroups samples not individuals. Therefore, for the rest of this article, we have $\hat{\theta}_i$ instead of X_i as this:

$$\left.\begin{aligned} Z_i &= \lambda_1\hat{\theta}_i + (1-\lambda_1)Z_{i-1} \\ W_i &= \lambda_2 Z_i + (1-\lambda_2)W_{i-1}, i = 1,2,\dots, \end{aligned}\right\}$$

The available DEWMA charts in the literature are mostly based on individual observations as discussed earlier. In this chapter, we shall explore the DEWMA chart with the five estimators defined earlier in chapter 2; evaluating the performance of the DEWMA based on these estimators. These estimators have been proved by Nazir *et al*[19] and Schoonhoven *et al*[17] to be robust both in phase I and phase II for classical CUSUM. The estimators are mean, median, midrange, TriMean and Hodges-Lehmann estimators.

Therefore, these estimators are then incorporated in the DEWMA charts to have our proposed charts. The mean is included in the estimators to serve as a yardstick with which we evaluate others since all others are unbiased estimator of mean. The robustness of these charts with respective estimators are thereby compared and analyzed in the upcoming section of this chapter.

3.4 PERFORMANCE EVALUATION OF THE PROPOSED CHARTS

In this study, we have considered the standardized normal distribution as the process environment. The performance measure is the run length, which is a random variable, representing the number of plotted statistics before a signal occurs. The average and standard deviation of this random variable are termed as average run length (ARL) and standard deviation run length (SDRL) respectively. With the aid of a Monte-Carlo simulation, we generated a 10^5 random numbers each of sample size $n = 5$, from a standard normal distribution with $\mu = 0$ and $\sigma^2 = 1$. We calculated each of the five estimators aforementioned, and independently used them in computing the DEWMA statistic.

The DEWMA statistic Z_i are then plotted against their control limits. Whenever Z_i is greater than its limits, the process is terminated and the point i is recorded. We repeat this algorithm for 10^5 runs. The average of all points i recorded for which the process sent a signal are then taken to be either in-control ARL's denoted as ARL_0 when there is no shift in the process mean or out-of-control ARL's denoted as ARL_1 when there exist a drift in the mean process. The shifts that we considered in this study lie in the range of $[0,1]$ both inclusive.

To accomplish one of the aims of this chapter, which is examining the robustness of the DEWMA charts, we have introduced some location and variance contaminations to the standard normal environment and studied their behaviors. Therefore, the simulation results to be presented will be in three-folds, the uncontaminated normal environment, the location

contaminated and variance contaminated environments and each of the contamination is of two cases, the 5% and 1% contaminations.

3.4.1 Uncontaminated Normal Environment

The Table 3.0 shows the ARL's and SDRL's of the DEWMA chart for all the five estimators under a standard normal environment whose observations were drawn from $N(\mu_0 = 0,\ \sigma_0^2 = 1)$. From the Table, we can see that the ARL_0's for all the five estimators are clustering around the target $ARL = 370$. This implies all the estimators are doing quite fine, since they are all unbiased estimators for mean. For the out-of-control ARL's denoted as ARL_1's, control charts are better off based on how small they are. That implies how fast the chart is in detecting shifts in the process. From the Table again when ($\delta \neq 0$), the mean also has the best performance as its values are the least of all for all the cases. All these insinuate the mean estimator is the best of all, under an uncontaminated normal environment.

Next to the mean in performance is the Hodges- Lehmann and the trimean estimators. These two estimators incorporated in the DEWMA chart have a close and tight range with the mean in fast detecting shifts in the process. Finally, we have the midrange and median in that hierarchy.

Table 3.0: ARL and SDRL Values for the DEWMA Based on Different estimators Under -Uncontaminated Normal Environment with $\lambda_1 = 0.13,\ \lambda_2 = 0.45$ and different L for a specified $ARL_0 = 370$

		δ											
$\widehat{\theta}$	*Values*	**0**	**0.25**	**0.33**	**0.4**	**0.48**	**0.55**	**0.63**	**0.7**	**0.78**	**0.85**	**0.93**	**1**
$\overline{x}$	***ARL***	369.46	21.725	12.734	9.1683	6.7763	5.4106	4.243	3.6272	3.0395	2.6863	2.3508	2.1062
	SDRL	368.90	17.233	9.3561	6.1726	4.3765	3.3539	2.5582	2.1384	1.7544	1.501	1.2624	1.0946
$\widetilde{x}$	***ARL***	369.12	29.797	17.872	12.737	9.2665	7.3048	5.7269	4.8882	4.064	3.0587	3.0587	2.7488
	SDRL	376.66	25.427	13.906	9.1228	6.3213	4.7093	3.587	2.9857	2.4383	2.0768	1.7654	1.543
MR	***ARL***	369.33	27.618	16.466	11.613	8.4252	6.6848	5.3733	4.4414	3.7946	3.2982	2.8269	2.5606
	SDRL	376.01	23.183	12.493	8.1934	5.6115	4.3059	3.304	2.7027	2.2215	1.9141	1.5651	1.4353
TM	***ARL***	375.02	23.804	14.302	10.126	7.342	5.9433	4.6783	3.9853	3.335	2.9094	2.5079	2.2615
	SDRL	384.91	19.375	10.640	6.8365	4.8205	3.7715	2.8347	2.3355	1.9041	1.6408	1.3732	1.2134
HL	***ARL***	374.13	23.042	13.749	9.629	7.1455	5.6758	4.5314	3.8108	3.2086	2.7908	2.4258	2.1934
	SDRL	372.34	19.206	10.071	6.5183	4.6292	3.5163	2.7284	2.244	1.8391	1.5651	1.3046	1.1703

3.4.2 Location Contaminated Normal Environment (LCNE)

Introducing a localized location contamination to the normal environment implies that we generate some sets of observations of $(\varphi)100\%$ from a $N(\mu_0 = 0,\ \sigma_0^2 = 1)$ and $(1-\varphi)100\%$ from a $N(\mu_0 + \ \omega\sigma_0, \sigma_0^2)$, where $-\infty < \omega < \infty$. This constitutes the whole data set we used in simulating and calculating our proposed DEWMA statistic with respective estimators. We set $\omega = 4$ and $\varphi = 0.05\ and\ 0.01$ to be the two levels of contamination. The two levels were chosen as low as those in order to observe the behaviors of the charts vividly and to have a substantial comparison and contrast. In Tables 3.2 and 3.3 we present the ARL's and $SDRL$'s simulation results of DEWMA control charts for all the five estimators for 5% and 1% level of contamination respectively.

It was to a surprise but not disappointing, to see the outcomes of the simulations as depicted in Tables 3.2 and 3.3. We were expecting values close to the target $ARL = 370$, while the highest value of the ARL's is just little above a hundred. It is not disappointing as we see ARL values increasing and approaching the target as we decreased the level of contamination. Then we deduced that the drastic fall in the values of ARL's is because the disturbance in the environment is location based at the same time the control charts are measuring and monitoring location parameter. Despite the poor performance so-to-say, we could manage to compare the estimators based on their results. "When the preferable is not available, the available becomes the preferable". The median estimator turns out to have the greatest value of in-control ARL and next to it is the trimean estimator, the rest estimators are so low in-value that are incomparable with the first two, the least of all is the midrange estimator followed by mean and Hodges-Lehmann estimators hierarchically. For the out-of-control ARL's it's another scenario. The estimators with least performance

in the ARL_0 values are taking the lead in the ARL_1's. the mean, midrange and Hodges-Lehmann estimators possess ARL_1's better than trimean and median estimators. The overall comparison, weighing the fact that a good estimators should take balance between its ARL_0 performance and the ability to detect a small drift in the out of control stage (i.e. small ARL_1). Relatively the trimean estimator has the best performance.

Table 3.1: ARL and SDRL Values for the DEWMA Based on Different estimators Under 5% LCNE with $\lambda_1 = 0.13,\ \lambda_2 = 0.45$ anddifferent L for a specified $ARL_0 = 370$

		δ										
$\widehat{\theta}$	*Values*	**0**	**0.25**	**0.33**	**0.4**	**0.48**	**0.55**	**0.63**	**0.7**	**0.78**	**0.85**	**0.93**	**1**
$\bar{x}$	***ARL***	24.157	7.7666	5.9536	4.9424	4.05	3.5127	3.0045	2.6206	2.3819	2.1372	1.9125	1.7654
	SDRL	23.397(	6.1887	4.3748	3.5635	2.7987	2.3576	1.9422	1.6276	1.4287	1.2354	1.0437	0.9294
$\tilde{x}$	***ARL***	119.77	17.472	11.911	9.1399	7.1342	5.8426	4.8583	4.1666	3.5279	3.1188	2.744	2.492
	SDRL	118.88	14.019	8.9528	6.5838	4.9066	3.7838	3.0546	25,967	2.1488	1.8394	1.5996	1.3912
MR	***ARL***	12.945	6.2977	5.3006	4.6226	4.0409	3.6271	3.1776	2.846	2.5676	2.3541	2.167	2.014
	SDRL	12.516	5.3819	4.354	3.5991	2.9804	2.624	2.222	1.9357	1.6566	1.4535	1.279	1.1413
TM	***ARL***	55.226	10.951	7.8844	6.3806	5.0255	4.2395	3.559	3.1279	2.6679	2.4041	2.1656	1.9653
	SDRL	53.834	8.5073	5.7959	4.408	3.3514	2.7684	2.23	1.919	1.5765	1.3845	1.1965	1.0605
HL	***ARL***	38.443	9.242	6.9741	5.689	4.4951	3.826	3.2369	2.802	2.4784	2.214	1.9943	1.8469
	SDRL	38.343	7.4099	5.1513	4.0234	3.0728	2.5466	2.0925	1.724	1.4987	1.2961	1.0999	0.9720

Table 3.2: ARL and SDRL Values for the DEWMA Based on Different estimators Under 1% LCNE with $\lambda_1 = 0.13,\ \lambda_2 = 0.45$ and different L for a specified $ARL_0 = 370$

		δ											
$\widehat{\theta}$	Values	0	0.25	0.33	0.4	0.48	0.55	0.63	0.7	0.78	0.85	0.93	1
$\bar{x}$	ARL	176.87	16.127	10.535	7.9238	5.9914	4.8551	3.9533	3.3745	2.8869	2.5476	2.2455	2.0111
	SDRL	179.67	12.754	7.6112	5.4702	4.0011	3.0691	2.4146	1.9927	1.6645	1.4237	1.2217	1.0599
$\tilde{x}$	ARL	319.19	26.690	16.645	11.758	8.7688	6.954	5.5786	4.6833	3.915	3.4499	2.9988	2.6859
	SDRL	328.73	22.744	12.501	8.4774	5.9904	4.5152	3.5373	2.8476	2.3285	2.0051	1.7169	1.5158
MR	ARL	90.581	16.438	11.618	9.0019	6.9546	5.8072	4.7281	4.0473	3.4609	3.0523	2.6663	2.3953
	SDRL	93.841	13.782	8.8977	6.6557	4.8531	3.8942	3.5433	2.5433	2.0944	1.8018	1.5173	1.3302
TM	ARL	263.29	19.461	12.014	8.8942	6.6251	5.4477	4.3505	3.6844	3.1247	2.7255	2.4218	2.1466
	SDRL	267.87	15.684	8.8158	6.2204	4.3982	3.3988	2.6693	2.2083	1.8212	1.56	1.329	1.1466
HL	ARL	185.42	16.561	10.678	7.8951	5.9858	4.8274	3.9669	3.361	2.8652	2.5139	2.2141	1.985
	SDRL	189.84	13.100	7.9151	5.6997	3.996	3.122	2.5067	2.0584	1.6736	1.4299	1.2256	1.0642

3.4.3 Variance-Contaminated Normal Environment (VCNE)

A variance-contaminated normal environment is a random variable generated from a distribution who has its dispersion disturbed. The environment is such that an $(\varphi)100\%$ observations are drawn from $N(\mu_0, \sigma_0^2)$ and complemented by $(1-\varphi)100\%$ observations from $N(\mu_0, \tau\sigma_0^2)$ where$-\infty < \tau < \infty$. Like we heard in the location-case, we are using two levels of variance contamination in this study, 5% and 1% respectively, $\tau = 9$ arbitrarily. The next two Tables 3.4 and 3.5 show the ARL and $SDRL$ results of DEWMA control charts with variance contaminated environments. On the other hand, we expect the simulation results to be better, compared with the location-contaminated case. The results are quite impressive, and the evaluation of the estimators' performances go the same way just as the location-contaminated environment.

Table 3.3: ARL and SDRL Values for the DEWMA Based on Different estimators Under 5% Variance-contaminated Normal Environment with $\lambda_1 = 0.13,\ \lambda_2 = 0.45$ anddifferent L for a specified $ARL_0 = 370$

		δ										
$\widehat{\theta}$	*Values*	**0**	**0.25**	**0.33**	**0.4**	**0.48**	**0.55**	**0.63**	**0.7**	**0.78**	**0.85**	**0.93**	**1**
$\bar{x}$	***ARL***	132.42	19.081	12.269	9.0424	6.78	5.4357	4.3441	3.7145	3.1031	2.7584	2.422	2.1516
	SDRL	138.26	16.286	9.602	6.8829	4.7675	3.7997	2.8825	2.4359	1.9385	1.684	1.4463	1.2349
$\tilde{x}$	***ARL***	269.72	28.756	17.578	12.603	9.1821	7.3127	5.7292	4.9007	4.0781	3.5668	3.0638	2.7856
	SDRL	276.43	24.652	13.886	9.4226	6.4126	4.8802	3.7014	3.0375	2.5106	2.1515	1.7716	1.5945
MR	***ARL***	58.569	20.189	14.235	10.728	8.0797	6.6556	5.3493	4.5655	3.843	3.3561	2.9299	2.6892
	SDRL	63.359	19.255	12.466	9.0452	6.4855	5.099	4.072	3.3325	2.6872	2.2967	1.9708	1.7824
TM	***ARL***	223.74	22.360	13.904	9.783	7.2389	5.8357	4.7173	3.9454	3.3096	2.8755	2.5313	2.2932
	SDRL	229.84	19.113	10.427	7.1373	4.9543	3.8899	2.9876	2.4664	1.9934	1.7008	1.4704	1.2804
HL	***ARL***	154.97	18.955	11.955	8.8033	6.5362	5.2557	4.1576	3.5678	3.0021	2.6666	2.3173	2.0905
	SDRL	160.95	16.116	9.4111	6.5351	4.5754	3.5275	2.7407	2.2609	1.8478	1.5865	1.3589	1.1687

Table 3.4: ARL and SDRL Values for the DEWMA Based on Different estimators Under 1% Variance-contaminated Normal Environment with $\lambda_1 = 0.13,\ \lambda_2 = 0.45$ anddifferent L for a specified $ARL_0 = 370$

		δ											
$\widehat{\theta}$	*Values*	**0**	**0.25**	**0.33**	**0.4**	**0.48**	**0.55**	**0.63**	**0.7**	**0.78**	**0.85**	**0.93**	**1**
$\bar{x}$	***ARL***	278.35	20.856	12.807	9.2794	6.7606	5.4039	4.3088	3.6518	3.0443	2.7119	2.3197	2.1083
	SDRL	283.93	17.152	9.489	4.4853	4.4853	3.4251	2.6703	2.1931	1.7385	1.5426	1.2783	1.1329
$\tilde{x}$	***ARL***	342.09	29.652	18.084	12.578	9.1323	7.2967	5.7781	4.8031	4.0362	3.531	3.0746	2.7361
	SDRL	347.53	25.608	14.153	9.1101	6.2515	4.7326	3.65	3.0007	2.4272	2.0596	1.7468	1.5421
MR	***ARL***	186.40	26.002	15.861	11.291	8.3479	6.6886	5.3521	4.5235	3.7923	3.2973	2.8531	2.576
	SDRL	194.20	22.207	12.428	8.2936	5.9306	4.5572	3.4767	2.8737	2.2902	1.9783	1.6744	1.4861
TM	***ARL***	335.26	23.523	14.155	10.093	7.3951	5.8306	4.6593	3.9212	3.3282	2.8823	2.5061	2.2544
	SDRL	339.70	19.040	10.755	6.9901	4.8153	3.7294	2.8656	2.3594	1.9159	1.6367	1.388	1.2222
HL	***ARL***	328.83	22.286	13.673	9.8143	7.0745	5.6408	4.5502	3.7752	3.1622	2.8112	2.4736	2.2173
	SDRL	327.18	18.032	10.181	6.7638	4.6464	3.5999	2.8052	2.2547	1.8678	1.611	1.3825	1.196

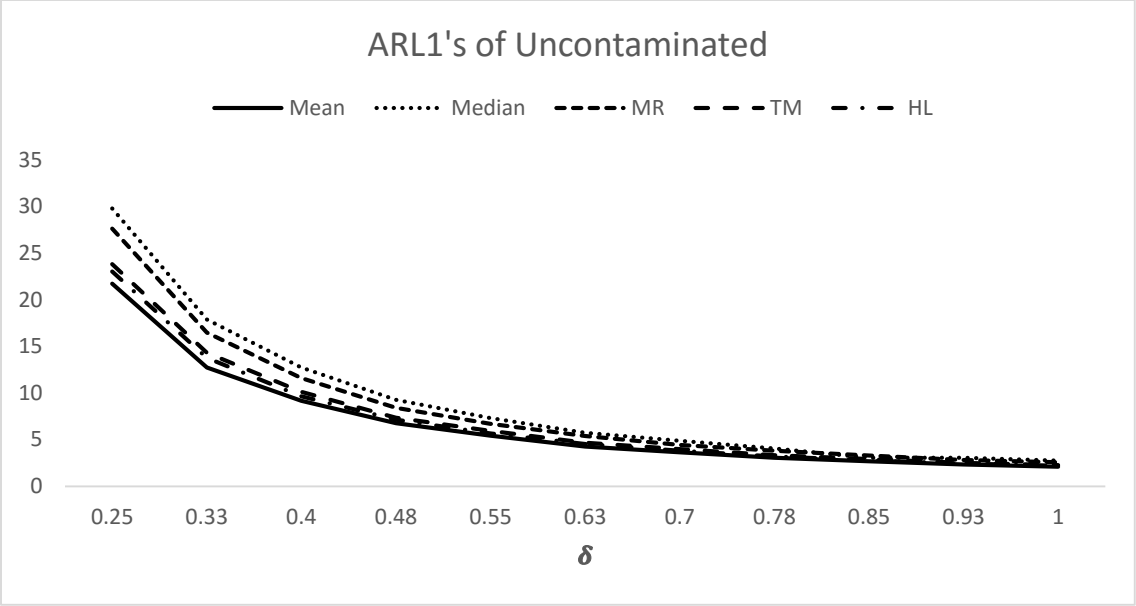

Figure 3.0*: The ARL_1's Curve of Uncontaminated DEWMA charts.*

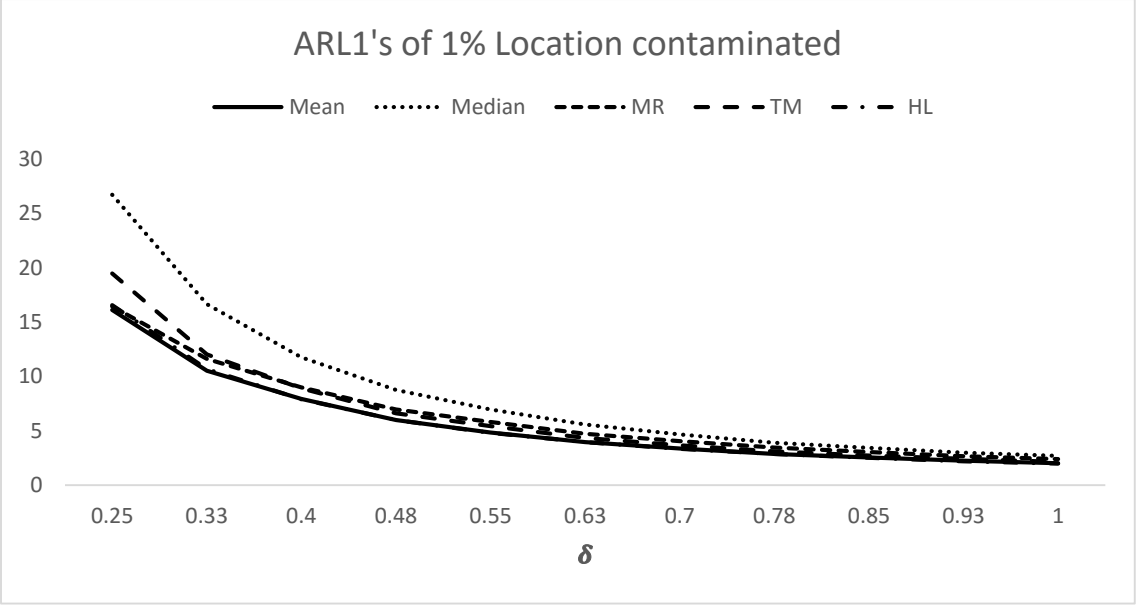

Figure 3.1*: The ARL_1's Curve of 1% Location Contaminated DEWMA charts.*

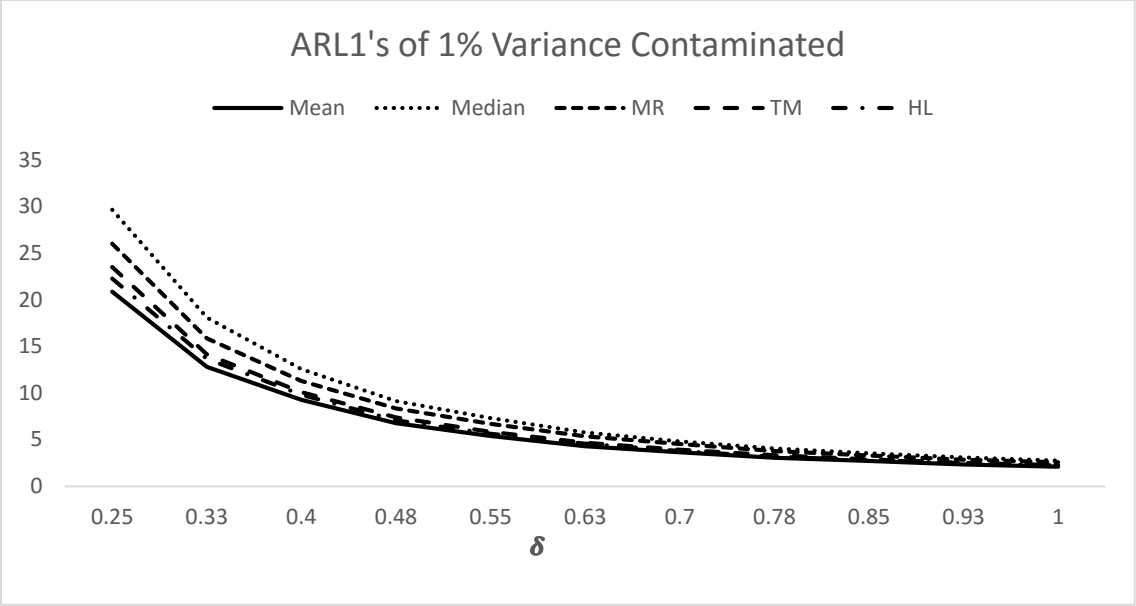

Figure 3.2*: The ARL_1's Curve of 1% Variance Contaminated DEWMA charts.*

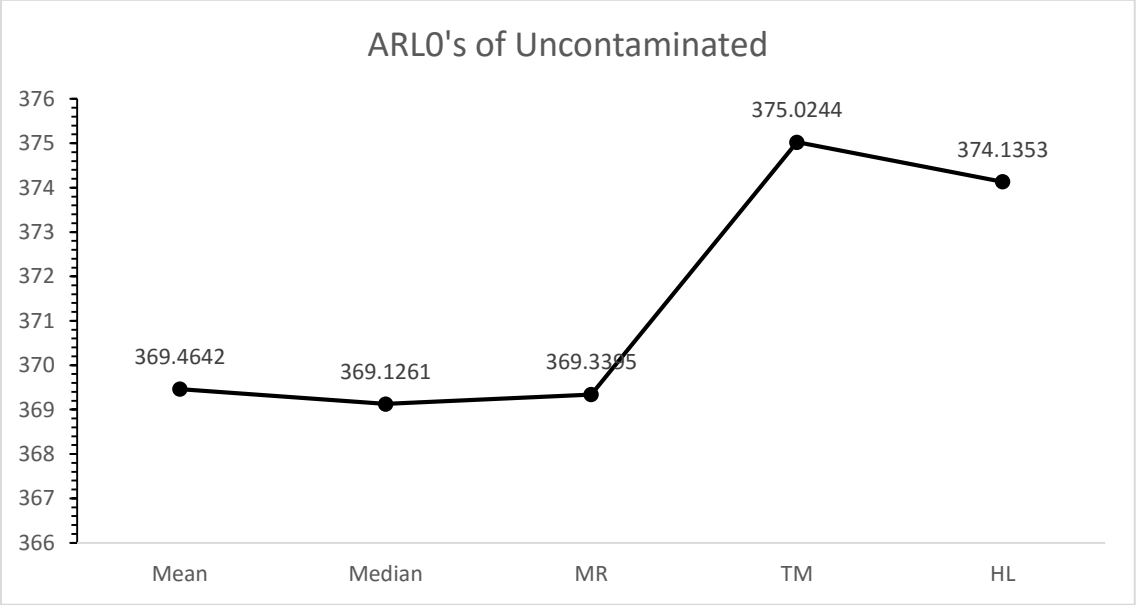

Figure 3.3*: The ARL_0's Values of Uncontaminated DEWMA charts.*

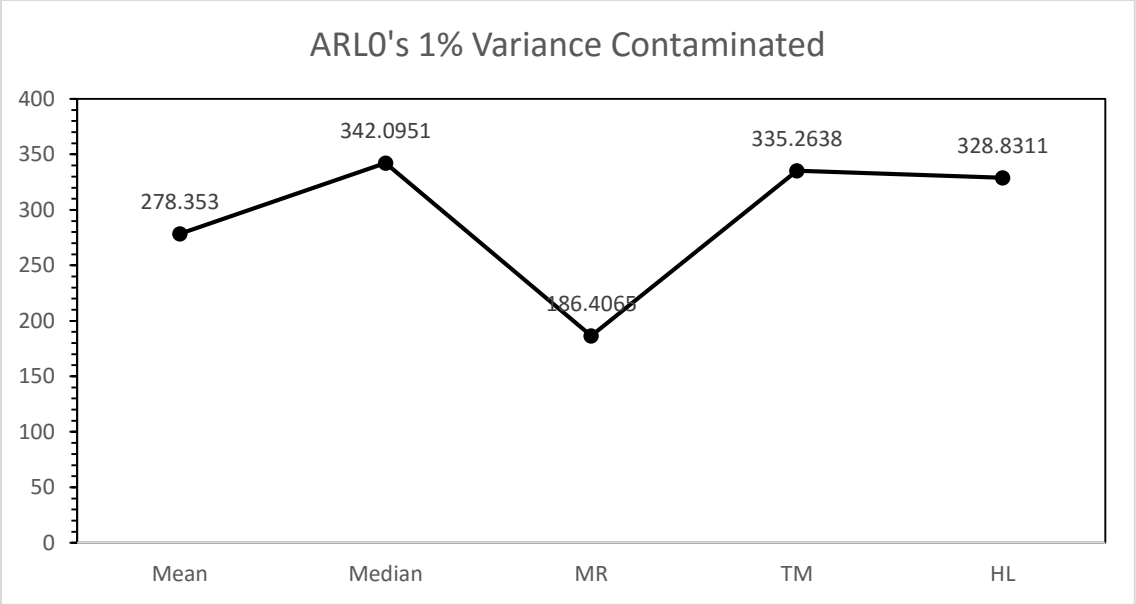

Figure 3.4*: The ARL_0's Values of 1% Location Contaminated DEWMA charts.*

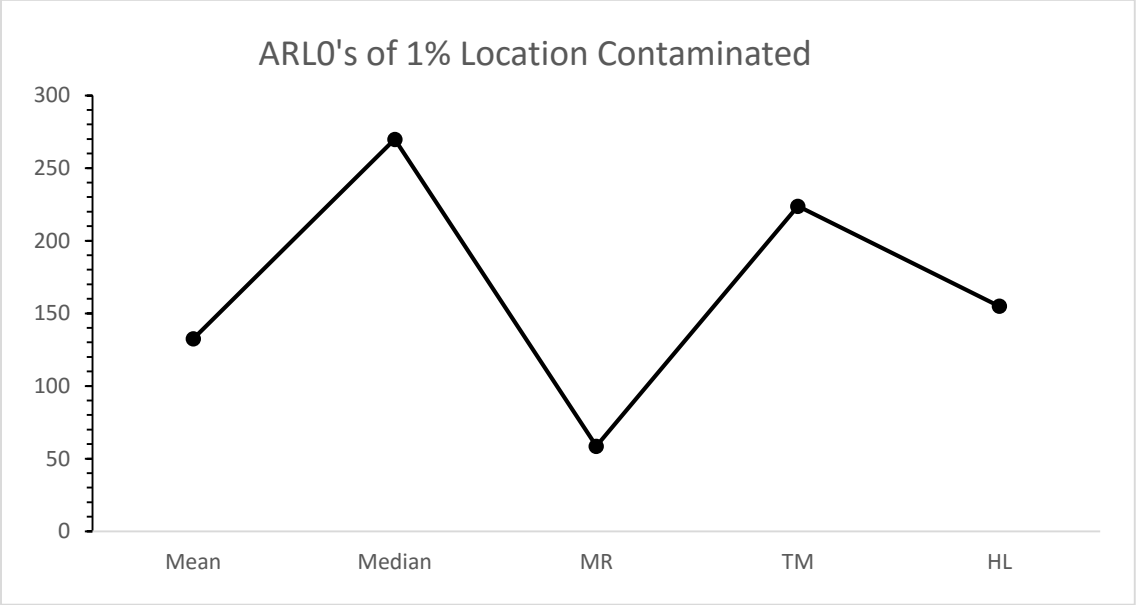

Figure 3.5*: The ARL_0's Values of 1% Variance Contaminated DEWMA charts.*

3.5 EFFECT OF PARAMETER ESTIMATION ON PROPOSED SCHEMES

In control charts, the parameters representing some quality characteristics to be measured in a process are often unknown, and therefore needs to be estimated. However, many research works published always assumed the process follows a particular probability distribution (mostly normal) and predict their parameters. This assumption makes the designation of the control charts in phase II stage easier, but in real life, the parameters are seldom known. In this study, we have investigated the effect of estimating the parameter being measured in our process on our proposed charts in the phase II analysis. The simulations result provided in section 4, were based on the fact that both the mean and variance of the process were known, but here, we estimated the location parameter mean assuming the variance is known. This estimation process will definitely cause some variability in the chart performance. Here, comes an expedient issue; selection of the best estimator for location parameter. Schoonhoven et al[17] has studied the robust location estimator in phase I, they inspected many estimators and concluded that the best and most robust against the localized and diffuse disturbances is the trimmed mean of the sample trimeans. The estimator is defined as follows:

$$\overline{TM}_{(\alpha)} = \frac{1}{k - 2\lceil k\alpha \rceil} \times \left[\sum_{v=\lceil k\alpha \rceil + 1}^{k-\lceil k\alpha \rceil} TM_{(v)} \right]$$

where $TM_{(v)}$ is the v^{th} ordered value of the sample trimeans. And $TM_i = \left(Q_{i,1} + 2Q_{i,2} + Q_{i,3}\right)/4$. where $Q_{i,q}$ is the q^{th} quartile of sample i, $q = 1,2,3$.

In our case, assuming that the spread parameter σ is known, and μ is unknown, we generated a preliminary sample of size $k = 50$ from which we calculated the 20% trimmed mean of sample trimeans (i.e. $\overline{TM}_{20}$) then the resulting estimate is used as our new location parameter. This is applied to all the five estimators in simulating the DEWMA control charts. The simulation is such that we generated our 5×10^5 of sample size 5 each with $\mu = \overline{TM}_{20}$ and $\sigma = 1$. In addition, all other computations follow. The same performance measures ARL and $SDRL$ of the simulations were adopted and recorded. Table 3.5 shows the ARL's and $SDRL$'s of the uncontaminated normal environment with an estimated location parameter. The variability is such that the in-control ARL's are a bit greater than the target $ARL_0 = 370$ and the out-of-control ARL's are much greater than expected as compared to the results in section 4.

This implies that the charts with estimated parameter has more longevity before raising a signal and a steady-slow way of sending alarms in the presence of little or small drifts in the process. All these are observed under the uncontaminated environment.

Just as the goal of any control chart, the effect of parameter estimation under the contaminated environments is awesome. The ARL_0's are much greater under parameter estimation in phase I as compared to when there is no parameter estimation and the ARL_1's are lower compared to their counterparts without the parameter been estimated preliminary. These effects were also felt in both the location and variance contaminations environments for all the estimators with exception of median estimator whose chart performance is conflicting with that of all other estimators. Astonishingly, the DEWMA chart based on median estimator with parameter estimation has its ARL's and SDRL's values lower than

otherwise, both in and out-of-controls ARL's respectively. Table 3.6 and 3.7 depict a better picture of the scenario.

Table 3.5: ARL and SDRL Values for the DEWMA Based on Different estimators Under -Uncontaminated Normal Environment with $\lambda_1 = 0.13,\ \lambda_2 = 0.45$ and different L for a specified $ARL_0 = 370$ with an estimated location parameter

		δ											
$\widehat{\theta}$	*Values*	**0**	**0.25**	**0.33**	**0.4**	**0.48**	**0.55**	**0.63**	**0.7**	**0.78**	**0.85**	**0.93**	**1**
$\bar{x}$	***ARL***	376.00	39.688	18.668	11.937	8.052	6.404	4.907	4.072	3.449	2.946	2.679	2.292
	SDRL	451.76	70.228	20.818	10.121	6.2064	4.1887	3.1734	2.4781	1.9548	1.6994	1.4745	1.2059
$\tilde{x}$	***ARL***	373.68	32.099	20.211	14.324	10.371	8.405	6.5561	5.5401	4.6674	4.1457	3.6349	3.3039
	SDRL	561.26	34.517	18.289	11.126	7.7713	5.8214	4.2408	3.2788	2.4582	2.0961	1.6775	1.422
MR	***ARL***	372.58	43.978	22.32	15.756	9.971	7.951	6.003	5.288	4.144	3.736	3.271	2.826
	SDRL	414.74	65.928	22.391	16.972	7.2518	5.8216	4.033	3.3323	2.4673	2.3134	1.9124	1.5447
TM	***ARL***	375.07	36.433	20.442	12.696	9.061	6.773	5.462	4.528	3.656	3.372	2.823	2.529
	SDRL	445.23	45.165	22.518	11.664	6.7111	4.808	3.5408	2.6916	2.061	1.9676	1.557	1.3947
HL	***ARL***	366.08	23.656	14.098	10.464	7.5642	6.2162	5.0961	4.4324	3.8326	3.451	3.1086	2.8492
	SDRL	640.74	25.337	12.557	8.012	5.0327	3.6554	2.5866	2.1049	1.6728	1.3715	1.1596	1.0304

Table 3.6: ARL and SDRL Values for the DEWMA Based on Different estimators Under 1% LCNE with $\lambda_1 = 0.13$, $\lambda_2 = 0.45$ and different L for a specified $ARL_0 = 370$ with an estimated location parameter

		δ											
$\widehat{\theta}$	Values	0	0.25	0.33	0.4	0.48	0.55	0.63	0.7	0.78	0.85	0.93	1
$\bar{x}$	ARL	197.85	6.239	3.81	3.018	2.241	1.94	1.553	1.4	1.255	1.137	1.074	1.053
	SDRL	26.573	4.4571	2.285	1.6627	1.1611	0.9947	0.6869	0.5953	0.4880	0.3746	0.2731	0.2241
$\tilde{x}$	ARL	196.75	8.4208	5.5135	4.1926	3.3015	2.8099	2.3912	2.1461	1.925	1.7667	1.6114	1.4886
	SDRL	286.70	5.9691	3.257	2.1445	1.4634	1.1331	0.862	0.7334	0.6392	0.5854	0.5535	0.5229
MR	ARL	102.65	7.019	4.579	3.601	2.656	2.32	1.885	1.597	1.415	1.286	1.206	1.108
	SDRL	124.57	4.9739	2.9258	2.2906	1.4247	1.1668	0.9574	0.7397	0.6043	0.5181	0.431	0.3201
TM	ARL	138.69	5.319	3.502	2.806	2.806	2.223	1.833	1.401	1.233	1.153	1.083	1.05
	SDRL	209.27	3.4396	2.0075	1.5681	1.5681	1.1307	0.7214	0.6038	0.4593	0.3711	0.2832	0.2181
HL	ARL	111.79	6.0023	4.2088	3.3723	2.7633	2.4193	2.1257	1.929	1.7454	1.6067	1.4733	1.3528
	SDRL	173.00	3.5976	2.0339	1.3873	1.0034	0.7989	0.6488	0.5791	0.5314	0.5263	0.5096	0.4805

Table 3.7: ARL and SDRL Values for the DEWMA Based on Different estimators Under 1% VCNE with $\lambda_1 = 0.13$, $\lambda_2 = 0.45$ anddifferent L for a specified $ARL_0 = 370$ with an estimated location parameter

		δ											
$\widehat{\theta}$	**Values**	**0**	**0.25**	**0.33**	**0.4**	**0.48**	**0.55**	**0.63**	**0.7**	**0.78**	**0.85**	**0.93**	**1**
$\bar{x}$	**ARL**	278.15(	6.832	4.309	3.2	2.445	1.963	1.629	1.43	1.278	1.17	1.092	1.054
	SDRL	351.06(	4.635	2.8099	1.9431	1.3419	1.0113	0.7536	0.6430	0.5069	0.4114	0.2926	0.2305
$\tilde{x}$	**ARL**	347.13(	9.826	6.0758	4.5265	3.4898	2.9154	2.5041	2.2191	1.9872	1.811	1.6449	1.5226
	SDRL	534.45(	7.2836	3.6933	2.4118	1.5762	1.1905	0.9380	0.7667	0.6529	0.6078	0.5539	0.5335
MR	**ARL**	185.19(	9.505	5.814	3.986	3.08	2.461	1.966	1.717	1.477	1.336	1.207	1.152
	SDRL	219.71(	7.9455	3.9409	2.4066	1.758	1.3514	1.0807	0.8412	0.6722	0.5526	0.4431	0.3861
TM	**ARL**	326.20(	8.086	5.005	3.575	2.636	2.216	1.794	1.543	1.298	1.232	1.133	1.081
	SDRL	391.68(	5.9306	3.0477	2.0274	1.4147	1.192	0.9002	0.7104	0.5229	0.4652	0.3397	0.273
HL	**ARL**	270.44(	6.8325	4.6548	3.6367	2.9135	2.5157	2.1997	1.9805	1.8018	1.6697	1.5034	1.3812
	SDRL	469.82(	4.3858	2.3547	1.5625	1.0919	0.8643	0.6769	0.5864	0.5469	0.5308	0.5173	0.4934

Conclusively, based on this study, we can confidently say the effect of parameter estimation on the phase II stage of control charting under the uncontaminated environment is not positive, whereas for the contaminated environments, the out-turn is quite impressive, as it increases the longevity of the chart before sending a signal. On the other hand, it speeds up the rate at which the chart detects small shifts in the process.

3.6 REAL-LIFE EXAMPLE

In this section we present the application of our proposed scheme on a real-life data set. The data set was a secondary data from a regional health care center, which is the waiting time (in minutes) of patients awaiting a colonoscopy procedure. It has 30 sample points with 5-sample size each, that is $n = 5$. The five proposed estimators were computed and used in constructing our DEWMA statistics. We introduced some disturbances to the original data set to envisage both the location and variance contaminated environments. The resulting control charts for the examples are shown below, part (a) of all the charts are the scatter plots of the sample points while parts (b, c, d, e and f) are for mean, median, midrange, trimean and Hodges-Lehmann estimators respectively.

Under the uncontaminated environments, all the estimators performed well as expected, since all the estimator are unbiased location estimators. As depicted in the Figures, none of the points in the chart crosses their respective limits.

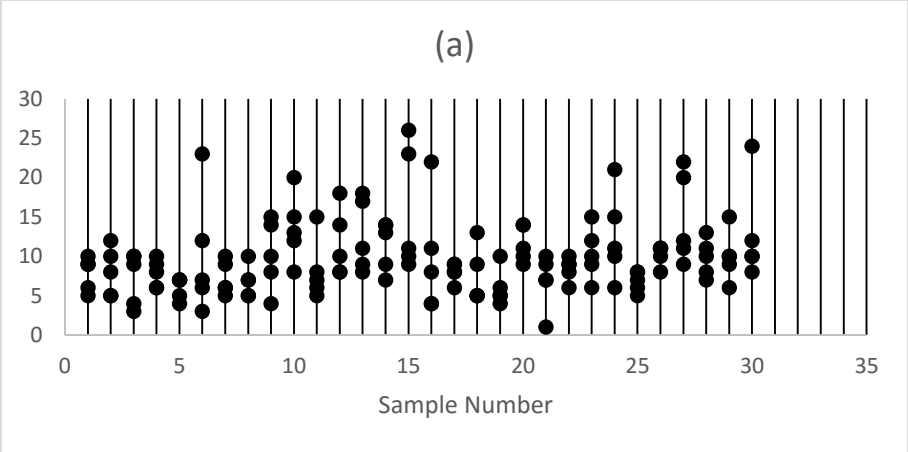

Figure 3.6: The scatter plot of the uncontaminated dataset

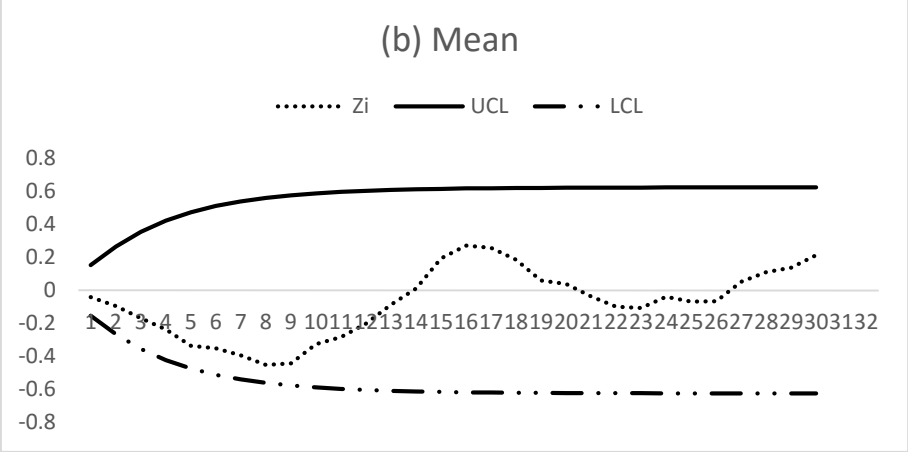

Figure 3.7*: Control Chart of mean DEWMA for uncontaminated dataset*

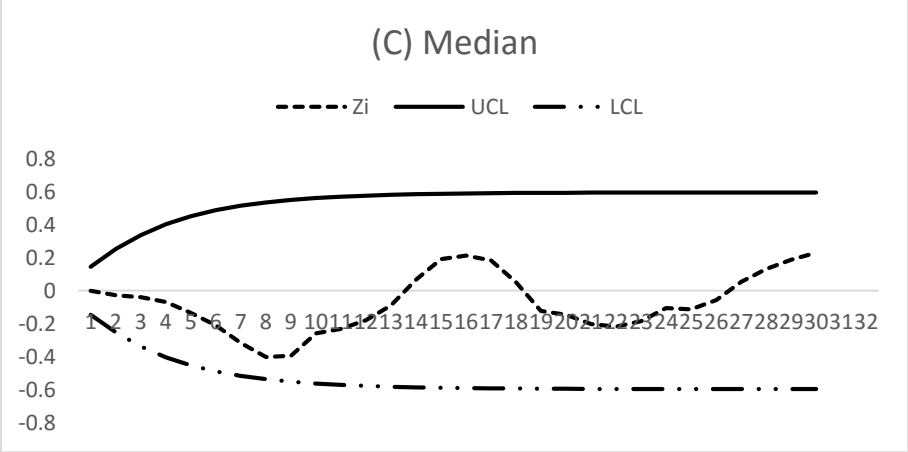

Figure 3.8*: Control Chart of median DEWMA for uncontaminated dataset*

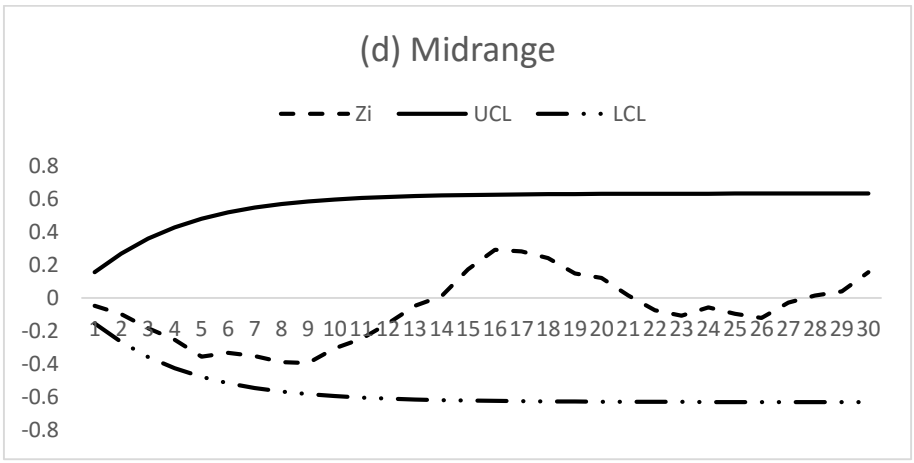

Figure 3.9*: Control Chart of Midrange DEWMA for uncontaminated dataset*

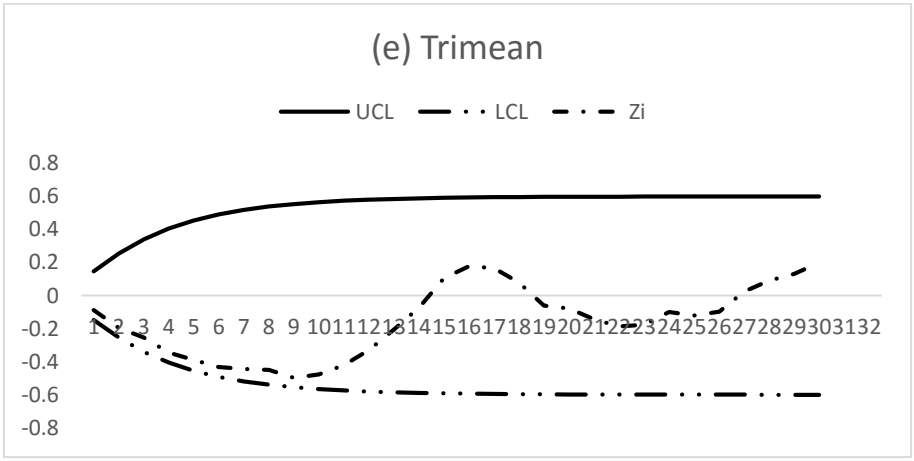

Figure 3.10*: Control Chart of trimean DEWMA for uncontaminated dataset*

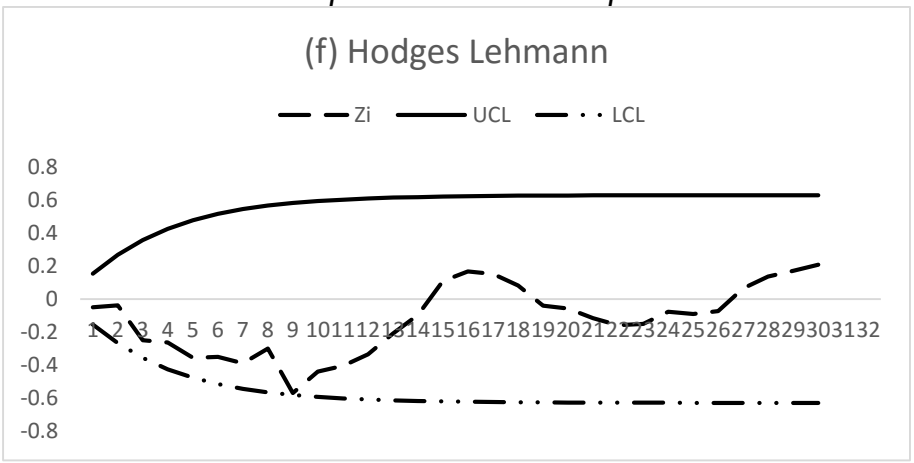

Figure **3.11***: Control Chart of HL DEWMA for uncontaminated dataset*

On the other hand, after introducing a 5% diffuse location contamination to the data set, the turn-out emulates the simulation results verbatim. As we can see in the respective simulation result, the midrange and mean estimators lost their performance efficiencies unlike the trimean, median and Hodges-Lehmann estimators which were able to retain their efficiencies despite the level of the contamination. The set of control charts below depict the behavior of each estimator for 5% location contamination, as for the mean in Figure 14, the chart is in out-of-control state for points 16 and 17 then returns to the in-control state. With respect to the midrange estimator, Figure 3.14, the chart was out of–control for at least five consecutive points starting from sample point 15 before regaining its stability. This concisely explains the weakness of these two estimators (mean and midrange) in a real life data set, while all other estimators performs honorably well even in the presence of contamination.

As for the variance contaminated environment, the scenario is not atypical of the previous. Though the level at which the midrange and mean estimators lose their efficiency here reduces as compared to the location contamination case, which is expected. Reason being that the estimators are location based and also measuring location parameter. Unlike the latter whose contamination is based on the variance parameter of the data set. The trimean estimator chart still emerged as the best control chart amongst its peer, next to it in performance are the median and Hodges-Lehmann estimators. The set of control chart below portrays this fact for the variance contamination environments.

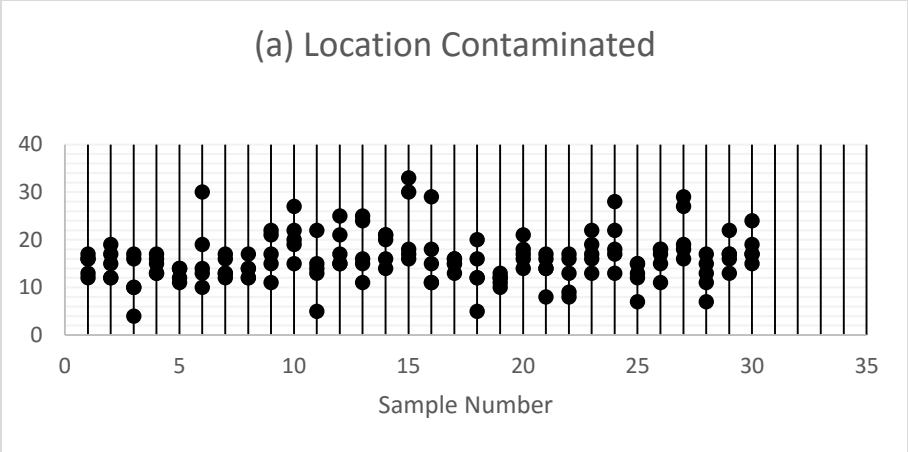

Figure3.12*: The scatter plot of the LCNE dataset*

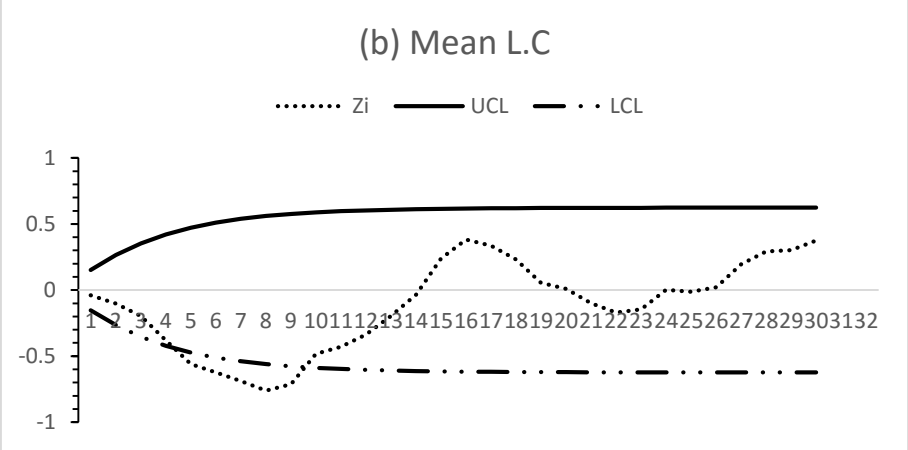

Figure 3.13*: Control Chart of mean DEWMA for LCNE dataset*

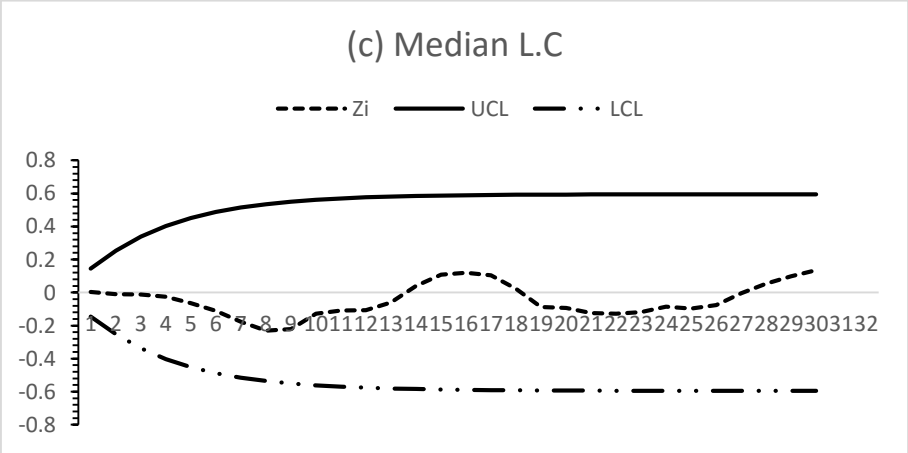

Figure 3.14: *Control Chart of median DEWMA for LCNE dataset*

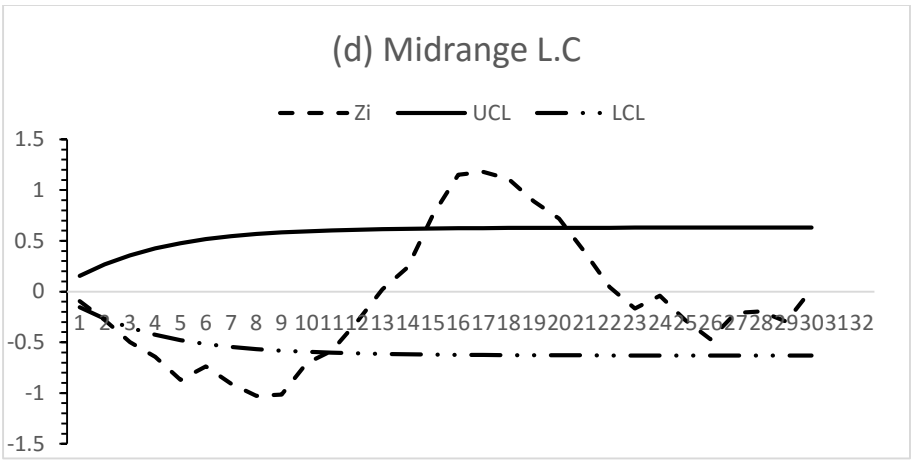

Figure 3.15*: Control Chart of mean DEWMA for LCNE dataset*

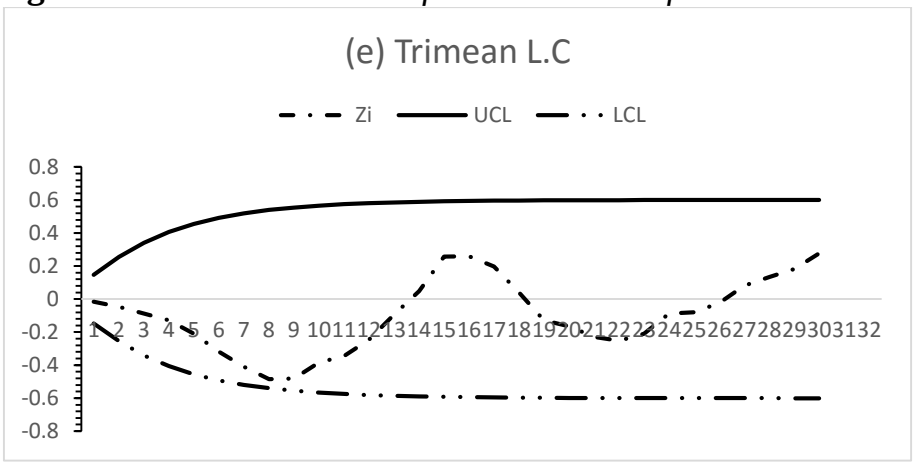

Figure 3.16: *Control Chart of trimean DEWMA for LCNE dataset*

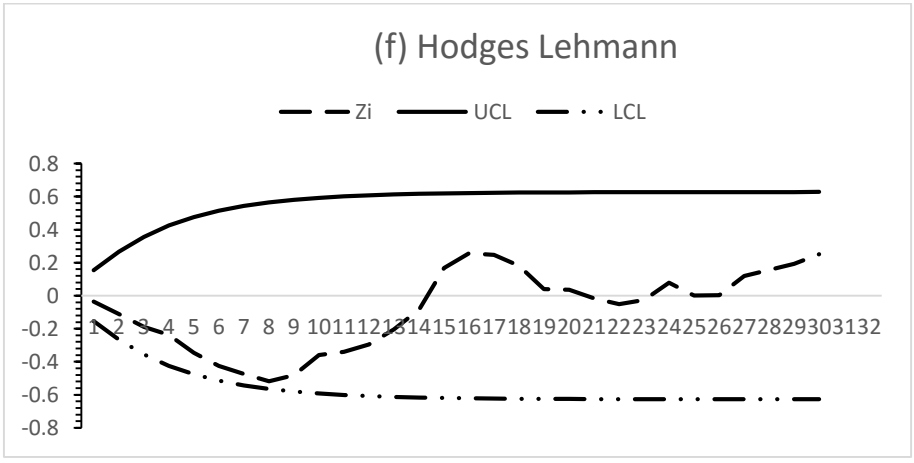

Figure 3.17: *Control Chart of HL DEWMA for LCNE dataset*

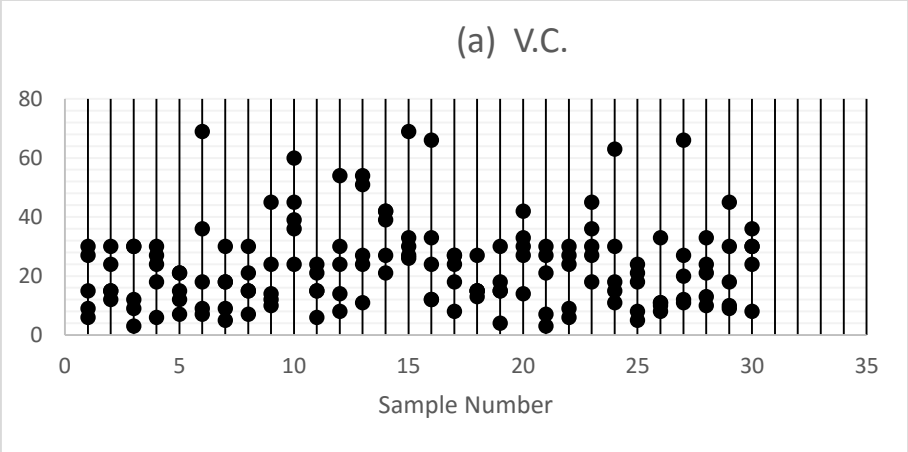

Figure3.18: *The scatter plot of the VCNE dataset*

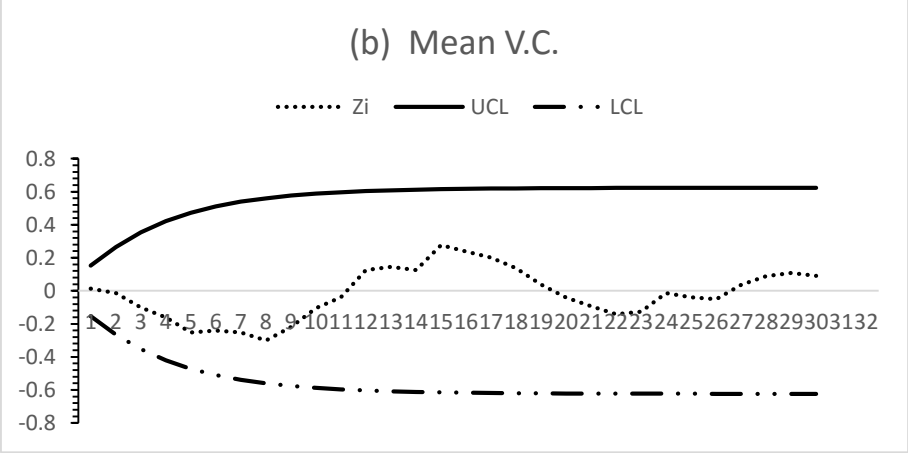

Figure 3.19*: Control Chart of mean DEWMA for VCNE dataset*

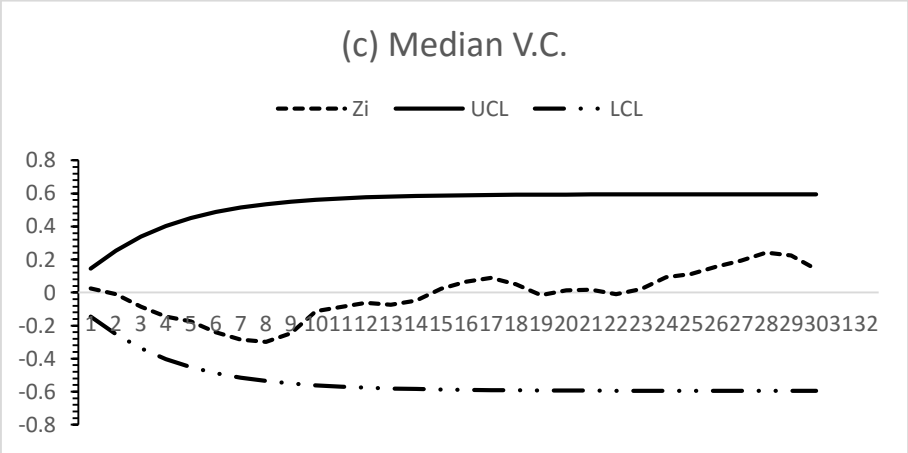

Figure 3.20*: Control Chart of median DEWMA for VCNE dataset*

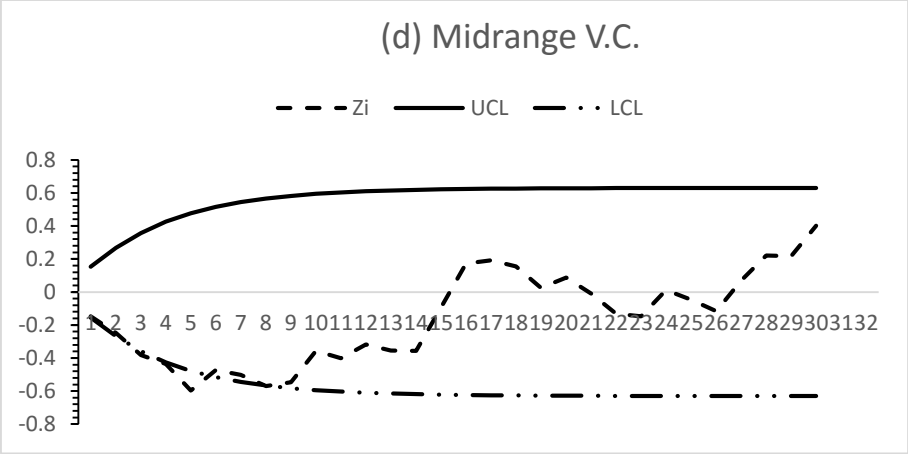

Figure 3.21*: Control Chart of midrange DEWMA for VCNE dataset*

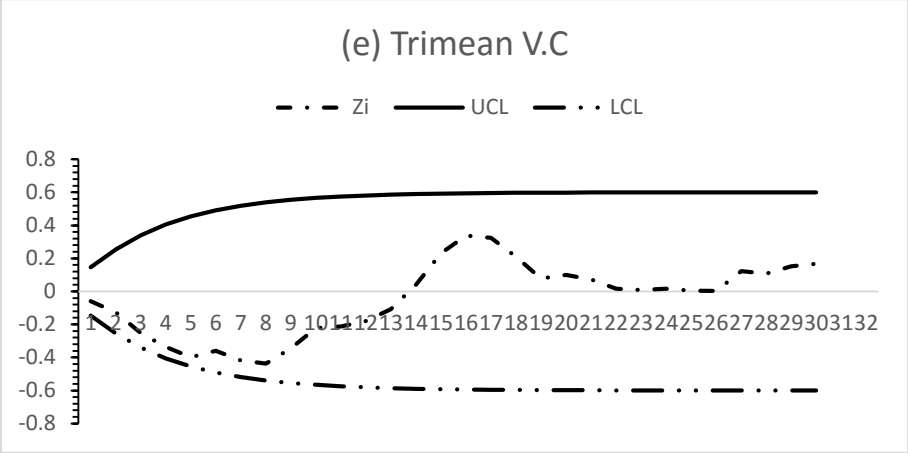

Figure 3.22: *Control Chart of trimean DEWMA for VCNE dataset*

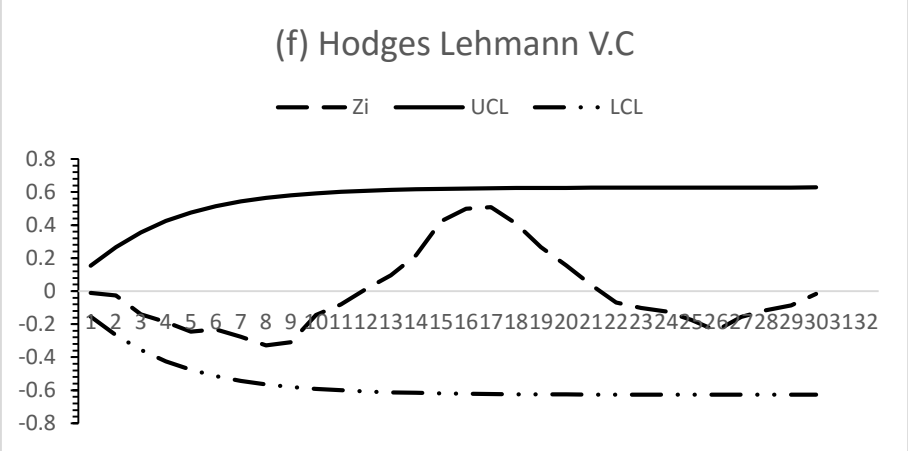

Figure 3.23: *Control Chart of HL. DEWMA for VCNE dataset*

3.7 SUMMARY AND CONCLUSION

It is truism that control chart is the best and most widely used of all the seven tools of statistical process control. This attribute of the control chart has fostered on its improvement, particularly the memory type. The major types of the memory charts are the Cumulative Sum (CUSUM) and the Exponentially Weighted Moving Average (EWMA). These two charts are comparatively competing on modifications by researchers. The EWMA on the other hand, has been extended to Double EWMA chart amongst many modifications, to enhance its efficiency and ability to fast detect small and moderate shifts. The superiority of the Double EWMA charts over the classical EWMA charts has been proven in literature. We investigated this scheme by introducing contaminations in the environments from which the synthetic data were generated for the simulations, with this we found that the usual mean estimator used in measuring location parameter in case of subgroup data was not robust. This really propelled us in exploring the scheme with some robust estimators, we therefore suggested another four location estimators keeping the mean estimator as yardstick measure. The findings of the study shows that the trimean estimator is the best of all, followed by the Hodges-Lehmann estimator. Though the median has a better performance in terms of the in-control ARL in all environments, but it is slow in detecting the presence of small shifts in the process, that is, its out-of-control ARL's are higher than others.

It is no doubt the mean estimator is a better option under uncontaminated environments, however, we recommend the adoption of the robust estimators for both contaminated and uncontaminated environments. We further initiated the investigation on the effect of parameter estimation in the phase II analysis; this reveals to us that, using robust estimator

also boosts the performance of these schemes. We used a 20% trimmed mean of trimeans to envisage the fact.

In conclusion, we completed the study by applying the scheme to real-life data set in a regional health center. Expectedly, the example buttressed the simulation results.

CHAPTER 4

EFFORTLESS HOTCHPOTCH OF EWMA & DUAL CUSUM CONTROL CHARTS

4.1 INTRODUCTION

Control charts are the most essential technique of all the Statistical Process Control (SPC) tools, because not only they are widely used in industries and getting premium attention in academia, but also they are phasing-out the vestiges of other tools of SPC. What control charts actually do is to design a retrospective study of a particular process, define its in-control state, assess the stability of the process over time and construct a reliable control limit for the process in Phase I, while in Phase II, it monitors and measures the process based on the designed limits therein to prevent/report any variation that could disturb the process from its target. The variations are of two types: Natural cause and assignable cause of variation. The presence of the former in a process doesn't take it from the fold of being in statistical control, whereas on the other hand, the latter purportedly puts a process in an out-of-control state.

The sensitivity of the memory chart has elevated its utilization in the recent time, and aroused the interest of researchers in looking at the modification of these charts in order to increase their sensitivity and detection ability. Scores of publications are available in literature on modifications of the Cumulative Sum (CUSUM) and Exponentially Weighted Moving Average (EWMA) charts; the two types of memory-less charts. Lucas[5], producing a chart that measures both large and moderate shifts in process, introduced the combination of the Shewhart and CUSUM statistics to form a single control chart. Lucas

and Saccucci in a similar way designed a Shewhart-EWMA chart to enhance the performance of the chart. Earlier before these, same Lucas and Crosier[7] initiated the Fast Initial Response (FIR) idea, incorporated in the CUSUM chart to produce a head start. Similarly, Steiner[29] introduced FIR to EWMA chart. In the name of increasing the sensitivity of the CUSUM chart, Yaschin[30] designed a weighted CUSUM that assigns different weights to past information. Another approach was that of Jiang et al[31] where he uses an adaptive CUSUM procedure with EWMA-based shift estimators in which a range of shifts is targeted. Riaz et al[16] improved the CUSUM control charts performance by applying the idea of run rules schemes. On the other hand, Abbas et al[32] did the same to EWMA chart.

Abbas et al[8] pioneered the mixture of two memory-type charts to increase its ability for detecting small and moderate shifts. He proposed a combination of EWMA and CUSUM structure to form a single chart, named Mixed EWMA-CUSUM (MEC) chart. In contrary, Zaman *et al*[33]. designed a Mixed CUSUM-EWMA (MCE) statistic. The difference between MEC and MCE will be explained in the next section.

Dualizing a one-sided CUSUM statistics was an idea of Zhao *et al*[12] that they tapped from modifying Lordens strategy of combining a set of infinite numbers of classical CUSUM and monitors them simultaneously. Zhao *et al.* instead made a finite number (in fact two) of combination, monitors them simultaneously and was able to produce a simulation results unlike Lordens. The Dual-CUSUM has since then being proved to out-perform the classical CUSUM. Series of articles are available for dualizing the EWMA statistic. Shamma and Shamma[23], Zhang[24], Mahmoud and Woodall[25] and Alkahtani [26], have all independently studied the Double EWMA chart, investigated the signal

resistance measure of it, emphasized its superiority overs the classical EWMA and evaluated its robustness against it.

In this chapter, we aimed at enhancing the performance of memory-type charts by incorporating both the mixture and dualization of these charts. We proposed a Mixed Dual CUSUM- EWMA chart for location process monitoring in order to improve the small-shift detection ability increasing the efficiency. We examined the scheme with different location estimators to inculcate robustness for withstanding different disturbances.

The remainder of this chapter is organized as follows; section 2 entails the explanation of the following charts: Mixed EWMA-CUSUM and Mixed CUSUM-EWMA. While section 3 chronicles the proposed scheme; Mixed Dual-CUSUM EWMA (MEDC), the performance evaluation of the chart is depicted in section 4. Section 5 reveals the comparison of the proposed chart with its existing counterparts. Conclusively, in section 6, we have the summary and conclusion.

4.2 Mixed EWMA-CUSUM & Mixed CUSUM-EWMA

What distinguishes the memory-type charts from the memory-less chart is that the former makes use of the past information of any process in addition to the present for designing its structure. This feature enables it to detect moderate and small shift in the process under consideration faster the latter. The two types of the memory charts are the Cumulative Sum (CUSUM) and Exponentially Weighted Moving Average (EWMA), which are well explained in earlier chapters. The possible mixtures of these two charts are discussed in this section.

4.2.1 THE MIXED EWMA-CUSUM (MEC) CHART

Abbas et al[8] pioneered the assortment of two memory-type charts. He proposed the combination to improve the sensitivity of the charts. The MEC chart was designed such that the EWMA statistic was embedded in the CUSUM statistic and the chart structure looks more like the CUSUM chart. The MEC chart is defined as follows:

$$\left.\begin{aligned} M_i^+ &= max[0, (Q_i - \mu_0) - a_i + M_{i-1}^+] \\ M_i^- &= max[0, -(Q_i - \mu_0) - a_i + M_{i-1}^-] \end{aligned}\right\}$$

Where M_i^+ and M_i^- are the upper and lower CUSUM statistics respectively which are initially set to zero (i.e. $M_0^+ = M_0^-$ =0), a_i is a time-varying reference value for the chart like k in the classical CUSUM and Q_i, the EWMA structure is defined as;

$$Q_i = \lambda_q Y_i + (1 - \lambda_q) Q_{i-1}$$

Where λ_q is the smoothing constant just like λ in the classical EWMA, such that $0 < \lambda_q < 1$ and the initial value of Q_i statistics is set equal to the target mean (i.e. $Q_0 = \mu_0$). Then the mean and variance of Q_i statistics are given as:

$$Mean(Q_i) = \mu_0, \text{ and } \quad Var(Q_i) = \sigma_Y^2 \left(\frac{\lambda}{(2-\lambda)}\left[1 - (1-\lambda)^{2i}\right]\right)$$

the Y_i is the i^{th} sample point of the process under consideration with individual observations. The two parameters of this chart like k and h in the classical CUSUM are both time-varying due to the effect of the variance of the EWMA structure therein, and they are defined as:

$a_i = a^* \sigma_Y \sqrt{\left(\frac{\lambda}{(2-\lambda)}[1-(1-\lambda)^{2i}]\right)}$, and $b_i = b^* \sigma_Y \sqrt{\left(\frac{\lambda}{(2-\lambda)}[1-(1-\lambda)^{2i}]\right)}$.

The constants a^* and b^* are chosen dependently for a specified ARL_0 value. The statistics M_i^+ and M_i^- are then plotted against b_i, at any point the former is greater than the limit b_i, the process is said to have shifted above the target value and whenever the latter is greater than b_i, the process has shifted below the target value. (See Abbas *at el*.)

4.2.2 THE MIXED CUSUM-EWMA (MCE) CHART.

The hotchpotch of the MCE scheme is a total reverse of the previous. Where the previous uses the EWMA structure as an embedment of the CUSUM statistics, the mixed CUSUM-EWMA chart was designed making the EWMA statistics the predominant part of the medley and making the CUSUM chart subsidiary. The godfather of this alternative idea was Zaman *et al*[33]. where he defined is proposed scheme with two statistics:

$$\left.\begin{aligned} MCE_t^+ &= (1-\lambda_c)MCE_{t-1}^+ + \lambda_c C_t^+ \\ MCE_t^- &= (1-\lambda_c)MCE_{t-1}^- + \lambda_c C_t^- \end{aligned}\right\}$$

Both C_t^+ and C_t^- are the classical CUSUM statistics as in equation (1). Similarly, λ_c lies strictly between 0 and 1(i.e.$0 < \lambda_c \leq 1$), it should be noted that the level of the sensitivity of this chart depends largely on the magnitude of the parameter λ_c and its equivalent in the EWMA related charts. The initial values of this schemes are take equal the target mean of C_t^+ and C_t^- respectively. Which implies that $MCE_0^+ = MCE_0^- = \mu_c$, and then mean and variance of the statistics are given as:

$$Mean(C_t^+) = Mean(C_t^-) = \mu_{C_t}$$

$$Var(C_t^+) = Var(C_t^-) = \sigma_{C_t}^2$$

Adopting these parameters, the control limit of MCE scheme is defined as

$$UCL_t = \mu_{C_t} + L_C\sigma_{C_t}\sqrt{\left(\frac{\lambda}{(2-\lambda)}[1-(1-\lambda)^{2t}]\right)}$$

where L_C is the width coefficient. Whenever MCE_t^+ gets beyond UCL_t at point t, the process is said to have shifted upwards and if MCE_t^- crosses the limit, the process is deemed as shifted downwards.

4.3 THE PROPOSED MIXED EWMA- DUAL CUSUM CONTROL CHART

This proposed chart comprises of a dual one-sided CUSUM statistics and a classical EWMA chart. The predominant part of the mixture is the dual-CUSUM (DCUSUM), as the EWMA structure is embedded in it. Before getting into details, it should be noted at this point that all charts aforementioned, were explained and designed for individual observation, and therefore can easily be extended for subgroup observation when sample size $n > 1$. The rest of the scheme will be based on subgroup observation.

The Dual-CUSUM designed by Zhao *et al*[12]. is being modified in this proposed scheme as:

$$\left.\begin{aligned} C_{1i}^+ &= \max[0, (Z_i - \mu_0) - K_1 + C_{1i-1}^+] \\ C_{2i}^+ &= \max[0, (Z_i - \mu_0) - K_2 + C_{2i-1}^+] \end{aligned}\right\}$$

The statistics are characterized by different parameters whose selection are based on some axioms, where Z_i is the EWMA structure defined as:

$$Z_i = \lambda\hat{\theta}_i + (1-\lambda)Z_{i-1}$$

and $\hat{\theta}_i$ is the estimator of i^{th} sample subgroup of the process under consideration. As usual, the smoothing factor of the EWMA structure lies $0 < \lambda \leq 1$ and the initial values of the statistics are made equal to zeros (i.e. $C_{10}^{+} = C_{20}^{+} = 0$) while the initial value of the EWMA structure is set equal to the target mean, that is, $Z_0 = \mu_0$. Logically, the mean and variance of Z_i are given as:

$$Mean(Z_i) = \mu_0\,,\ \ Var(Z_i) = \sigma_\theta^2\left(\frac{\lambda}{(2-\lambda)}\left[1-(1-\lambda)^{2i}\right]\right)$$

The parameters K_1, K_2 and H_1, H_2 are dependently chosen for the statistics C_{1i}^{+}, C_{2i}^{+} with a specified ARL_0 based on the following conditions:

$$\left.\begin{array}{c} K_1 < K_2 \\ H_1 > H_2 + K_2 - K_1 \\ {}^{K_2}/_{K_1} \cong {}^{H_1}/_{H_2} \end{array}\right\}$$

In order to avoid the redundancy and ineffeciancy of any of the two statistics, these three conditions must be met. These parameters are then used to construct the control limits as defined below;

$$UCL_t = \mu_0 + H_t\,\sigma_\theta^2\sqrt{\left(\frac{\lambda}{(2-\lambda)}\left[1-(1-\lambda)^{2i}\right]\right)}, whre\ t = 1{,}2.$$

The statistics are then plotted against their respective limits, at any point either of the two crosses its respective limit, the process is said to out of control.

The $\hat{\theta}_i$ can then be replaced by any of our five estimators of location parameters, which are: mean, median, midrange, trimean and Hodges-Lehmann estimators. The mean is kept as yardstick to measure and evaluate other estimators. The efficiency and robustness of this scheme is a factor of the estimator with which it was constructed. These five estimators have been examined by researchers and were proved to be able to withstand disturbances, both in the location and dispersions parameters.

4.4 PERFORMANCE EVALUATION OF THE PROPOSED SCHEME

The common and most widely used measure of control charts is the run length behavior of the scheme. The run length can be simply defined as a random variable representing the number of plotted statistics until a signal occurs. The average of this random variable is called Average Run Length (ARL), while its standard deviation is termed Standard Deviation of Run Length (SDRL). The ARL and SDRL are of two versions; the in-control denoted as ARL_0 and $SDRL_0$, that is, the ARL and SDRL of the chart when there is no shift in the process ($\delta = 0$) and the other one is the out-of-control ARL and SDRL, denoted as ARL_1 and $SDRL_1$, these are the ARL's and SDRL's when the process has shifted from its target mean ($\delta \neq 0$). The run length distribution of a control chart can be investigated via different techniques such as Markov chains, integral equations, Monte Carlo simulations and different types of approximations. For the course of this study, we've employed Monte Carlo simulation by developing an algorithm in Matlab, which records the run length, this algorithm was repeated 10^5 times, then the average and standard deviation of this random variable were calculated.

The algorithm for simulating the ARL's and SDRL's was that, we generated a 5×10^5 random numbers of sample size $n = 5$ each, from a standard normal distribution, calculated the five estimators explained above, standardized them (haven known the standard deviation of each estimator) and use them for constructing the mixed DCUSUM-EWMA chart for a range of shifts, combinations of K_1, K_2 and H_1, H_2 for a specified $ARL_0 = 370$ and different cases of λ. Recall that in the classical EWMA, if the smoothing factor $\lambda = 1$, the chart becomes a Shewhart chart (a memory-less chart) also if $\lambda = 1$ in the MEC chart, it becomes a classical CUSUM chart. Similarly, when $\lambda = 1$ in our proposed scheme, it results to the DCUSUM chart. We carried out this algorithm for different values of λ under the contaminated and uncontaminated normal environments.

4.4.1 Uncontaminated Normal Environment

By uncontaminated normal environment, we mean the standardized normal distribution, that is the random numbers used for the simulation were drawn from a $N(\mu_0 = 0, \sigma_0^2 = 1)$. Adopting the algorithm described above, we computed the ARL's and SDRL's of MEDC schemes for different values of EWMA structure smoothing factorλ, ($\lambda = 0.05, 0.25, 0.5 \; and \; 0.75$). we found out that the simulation result of MEDC scheme for $\lambda = 1$, is exactly the same as the DCUSUM scheme. Tables 4.1-4.4 show the result for uncontaminated environment for a specified $ARL_0 = 370$. When $\delta = 0$, the ARL_0's of all the five estimators are close to zero in all cases of λ. Though the scope of this study is not to compare the choice of λ, yet there is a room for that. Under this environment, the performance of the estimators are close because they are all unbiased estimator of the mean. Despite the closeness, the mean estimator distinguishes itself from others by having the

lowest out of control ARL's, followed by it are Trimean and Hodges-Lehmann estimators. Finally comes the Midrange and Median estimators in that hierarchy.

Table 4.0: ARL and SDRL Values for the MEDC chart Based on Different estimators Under -Uncontaminated Normal Environment with $k_1 = 0.125,\ k_2 = 0.5, \lambda = 0.05$ and different H_1&H_2 for a specified $ARL_0 = 370$

		δ											
$\widehat{\theta}$	*Values*	**0**	**0.25**	**0.33**	**0.4**	**0.48**	**0.55**	**0.63**	**0.7**	**0.78**	**0.85**	**0.93**	**1**
$\bar{x}$	***ARL***	366.089	24.4801	19.4975	16.7427	14.4539	13.0629	11.8424	10.9301	10.0706	9.4906	8.8767	8.4467
	SDRL	344.539	8.2732	5.3781	3.9756	3.0305	2.4843	2.067	1.7771	1.5211	1.3485	1.209	1.1067
$\tilde{x}$	***ARL***	372.624	28.4225	22.2724	19.029	16.2983	14.7037	13.2755	12.2472	11.2586	10.6112	9.9061	9.4069
	SDRL	353.26	11.1006	7.1213	5.2783	3.949	3.1957	2.6483	2.2744	1.9553	1.7486	1.5285	1.385
MR	***ARL***	374.758	28.0157	22.2052	18.9594	16.3649	14.7327	13.3478	12.2928	11.3892	10.6977	9.9824	9.4822
	SDRL	351.209	10.3085	6.7131	4.8665	3.703	3.0355	2.4866	2.142	1.8584	1.6477	1.4435	1.3098
TM	***ARL***	373.095	25.5795	20.2471	17.4403	14.9702	13.5185	12.2613	11.2598	10.4147	9.8149	9.1413	8.6851
	SDRL	351.153	9.0023	5.8263	4.3435	3.2708	2.6672	2.2202	1.9055	1.6278	1.4483	1.2884	1.1615
HL	***ARL***	370.133	25.1694	20.019	17.1855	14.8148	13.3889	12.1541	11.184	10.3537	9.744	9.078	8.6324
	SDRL	347.031	8.6743	5.6336	4.1698	3.1704	2.5976	2.1448	1.8503	1.5946	1.428	1.2536	1.1345

Table 4.1: ARL and SDRL Values for the MEDC chart Based on Different estimators Under -Uncontaminated Normal Environment with $k_1 = 0.125,\ k_2 = 0.5, \lambda = 0.25$ and different H_1&H_2 for a specified $ARL_0 = 370$

		δ											
$\widehat{\theta}$	*Values*	**0**	**0.25**	**0.33**	**0.4**	**0.48**	**0.55**	**0.63**	**0.7**	**0.78**	**0.85**	**0.93**	**1**
$\bar{x}$	***ARL***	375.516	20.382	14.6448	12.0808	10.1122	8.9164	7.9379	7.2518	6.6612	6.2291	5.8095	5.4801
	SDRL	358.902	10.1193	5.8632	4.0516	2.987	2.3501	1.8253	1.5512	1.3345	1.1609	1.0121	0.9155
$\tilde{x}$	***ARL***	373.819	25.2571	17.8416	14.4359	11.8675	10.3566	9.1401	8.3258	7.5863	7.0651	6.5361	6.1902
	SDRL	353.383	14.2354	8.4215	5.8705	4.08	3.1937	2.5129	2.0875	1.7492	1.5163	1.3033	1.1921
MR	***ARL***	373.533	23.6436	17.0487	13.9413	11.5689	10.1997	8.9863	8.1573	7.4596	6.9556	6.4971	6.151
	SDRL	360.143	12.8546	7.4109	5.3523	3.7804	2.9231	2.3355	1.9156	1.6126	1.3992	1.2433	1.1132
TM	***ARL***	370.105	21.2484	15.4792	12.5904	10.4067	9.2162	8.2037	7.4852	6.8253	6.3995	5.9449	5.6314
	SDRL	357.778	10.9963	6.5634	4.5181	3.1919	2.4989	2.0024	1.6771	1.4237	1.2428	1.0961	0.9896
HL	***ARL***	366.014	20.7726	15.0357	12.3588	10.3451	9.1785	8.1203	7.4307	6.8165	6.3443	5.922	5.6127
	SDRL	350.231	10.5867	6.1872	4.2853	3.0697	2.4531	1.9368	1.6267	1.389	1.2	1.0433	0.9636

Table 4.2: ARL and SDRL Values for the MEDC chart Based on Different estimators Under -Uncontaminated Normal Environment with $k_1 = 0.125,\ k_2 = 0.5, \lambda = 0.5$ and different H_1&H_2 for a specified $ARL_0 = 370$

		δ											
$\widehat{\theta}$	**Values**	**0**	**0.25**	**0.33**	**0.4**	**0.48**	**0.55**	**0.63**	**0.7**	**0.78**	**0.85**	**0.93**	**1**
$\bar{x}$	***ARL***	372.876	19.7464	13.5475	10.584	8.4904	7.2721	6.3062	5.6661	5.1366	4.7326	4.3663	4.111
	SDRL	363.623	11.9448	7.0122	4.6971	3.2683	2.4928	19.275	1.6117	1.3312	1.1327	0.9772	0.8879
$\tilde{x}$	***ARL***	375.338	25.6783	16.9272	13.1628	10.4076	8.864	7.5585	6.7581	6.0207	5.5181	5.0777	4.7401
	SDRL	369.646	18.0458	9.808	6.8138	4.739	3.6398	2.7401	2.2453	1.8127	1.5525	1.3316	1.1778
MR	***ARL***	374.672	24.1234	16.2008	12.6212	9.9972	8.5504	7.3043	6.5437	5.8739	5.4015	4.9717	4.6556
	SDRL	365.537	15.8469	8.9406	6.2109	4.2687	3.2536	2.4577	2.0391	1.6724	1.4208	1.2346	1.0908
TM	***ARL***	377.279	21.0826	14.3998	11.1258	8.9214	8.4603	6.5489	5.8816	5.3151	4.8965	4.5025	4.2577
	SDRL	372.828	13.5796	7.7353	5.2003	3.6123	3.3012	2.0817	1.7363	1.4382	1.2312	1.0803	0.9627
HL	***ARL***	372.214	20.4693	14.1004	10.9232	8.772	7.5093	6.5371	5.8766	5.3024	4.8894	4.5193	4.2358
	SDRL	367.654	13.0044	7.2028	4.9023	3.4148	2.6353	2.0419	1.6712	1.3927	1.1955	1.0541	0.9259

Table 4.3: ARL and SDRL Values for the MEDC chart Based on Different estimators Under -Uncontaminated Normal Environment with $k_1 = 0.125,\ k_2 = 0.5, \lambda = 0.75$ and different H_1&H_2 for a specified $ARL_0 = 370$

		δ											
$\widehat{\theta}$	Values	0	0.25	0.33	0.4	0.48	0.55	0.63	0.7	0.78	0.85	0.93	1
$\bar{x}$	ARL	372.488	20.9797	13.2315	10.074	7.87	6.6608	5.6323	5.0013	4.4103	4.0225	3.6815	3.4201
	SDRL	368.832	14.25	7.771	5.1787	3.6335	2.7881	2.0932	1.736	1.3984	1.1994	1.0321	0.89448
$\tilde{x}$	ARL	367.051	26.7022	17.6024	13.1064	9.9984	8.2701	6.8874	6.0281	5.3296	4.8056	4.3339	4.0247
	SDRL	357.116	19.0996	11.6561	7.7757	5.3863	3.9668	3.0187	2.4147	1.9899	1.6643	1.4117	1.2427
MR	ARL	375.118	25.4384	16.5699	12.3341	9.4653	7.8147	6.6573	5.8344	5.1472	4.6607	4.248	3.9529
	SDRL	367.521	17.7823	10.4617	6.9909	4.7409	3.5648	2.7435	2.1858	1.7971	1.5329	1.2943	1.1385
TM	ARL	367.906	22.1578	14.3277	10.7617	8.3229	6.9161	5.8487	5.1863	4.5878	4.1527	3.785	3.5416
	SDRL	351.877	15.6171	8.7933	5.983	4.0675	3.0178	2.3148	1.9108	1.5404	1.2954	1.1073	1.0188
HL	ARL	371.500	21.4299	14.0881	10.5707	8.21	6.8852	5.8255	5.1462	4.542	4.158	3.8058	3.5466
	SDRL	367.237	14.5863	8.5504	5.6242	3.7801	2.9104	2.1924	1.814	1.4707	1.2325	1.0883	0.9619

4.4.2 Location – Contaminated Normal Environment (LCNE)

A random variable X is said to be drawn from a LCNE, if it has $(\varphi)100\%$ observations from $N(\mu_0, \sigma_0^2)$ and $(1-\varphi)100\%$ observations from $N(\mu_0 + \omega\sigma_0, \sigma_0^2)$ where $-\infty < \omega < \infty$. The same we did for earlier chapters, we do for the MEDC scheme, to investigate how long the estimators could retain their performance in the presence of disturbances and violation of normality assumptions. This feature is the yardstick we measure the robustness of the scheme via the respective estimators. Generating the random variables from a 5% and 1% location contaminated environment with $\omega = 4$, goes a long way to reveal the robustness of each estimator. We shall limit the discussion of the simulation results on the 1% contamination only. The ARL_0's of all the estimators are lesser than 70% of the expected value 370. The contamination is the major factor for this gross reduction, added to it is the fact that its location based, while the scheme is also monitoring the location parameter.

However, the aim of the study is not defeated; we noticed that both midrange and mean turn worst as their performance are not comparable with other estimators. Whereas, the median, trimean and Hodges-Lehmann estimators are relatively performing better. Though the median has the highest ARL_0, but its inability to send quick signals in the presence of small shift as revealed by smaller values of ARL_1's deprived it from being the best estimators. On the other hand, trimean and Hodges-Lehmann combined the features of being robust against contamination and the ability to send signals quickly in the presence of any shift. Relatively, the trimean is the overall best estimator, followed by Hodges-Lehmann and median estimators respectively under the location contaminated environment.

Table 4.4: ARL and SDRL Values for the MEDC chart Based on Different estimators Under LCNE with $k_1 = 0.125,\ k_2 = 0.5, \lambda = 0.05$ and different H_1&H_2 for a specified $ARL_0 = 370$

		δ											
$\widehat{\theta}$	Values	0	0.25	0.33	0.4	0.48	0.55	0.63	0.7	0.78	0.85	0.93	1
$\bar{x}$	ARL	133.84	21.818	17.917	15.638	13.743	12.458	11.334	10.514	9.7911	9.1896	8.6413	8.2233
	SDRL	112.20	7.1546	4.9027	3.7515	2.9361	2.4227	2.0011	1.7368	1.534	1.3687	1.2261	1.1266
$\tilde{x}$	ARL	255.57	27.127	21.631	18.484	16.040	14.427	13.065	12.050	11.158	10.447	9.7987	9.3106
	SDRL	236.95	10.502	18.484	5.0449	3.885	3.1977	2.5791	2.1829	1.9335	1.706	1.5328	1.3963
MR	ARL	90.226	23.034	19.113	16.782	14.849	13.596	12.405	11.546	10.727	10.120	9.5138	9.0801
	SDRL	71.245	8.1049	5.6801	4.4273	3.5576	3.0247	2.498	2.2233	1.8963	1.7072	1.5134	1.4123
TM	ARL	201.82	23.802	19.106	16.579	14.496	13.126	11.864	11.016	10.205	9.608	8.9983	8.5498
	SDRL	181.38	8.0302	5.3102	4.0567	3.1318	2.6105	2.1455	1.8475	1.6191	1.4292	1.2724	1.1545
HL	ARL	186.97	23.349	18.941	16.410	14.322	12.973	11.786	10.931	10.151	9.5285	8.9661	8.5095
	SDRL	165.46	7.8823	5.2046	3.9304	3.0591	2.514	2.0884	1.8274	1.5873	1.3986	1.2691	1.1414

Table 4.5: ARL and SDRL Values for the MEDC chart Based on Different estimators Under LCNE with $k_1 = 0.125,\ k_2 = 0.5, \lambda = 0.25$ and different H_1&H_2 for a specified $ARL_0 = 370$

		δ											
$\hat{\theta}$	Values	0	0.25	0.33	0.4	0.48	0.55	0.63	0.7	0.78	0.85	0.93	1
$\bar{x}$	ARL	135.27	17.129	13.121	11.046	9.4284	8.4286	7.6228	6.9789	6.4151	6.0317	5.6334	5.3166
	SDRL	119.81	8.1098	5.1531	3.7142	2.7843	2.1765	1.8144	1.52	1.3057	1.1527	1.0264	0.9126
$\tilde{x}$	ARL	267.04	23.284	17.008	13.888	11.467	10.092	8.9242	8.1543	7.4562	6.9739	6.4576	6.1236
	SDRL	255.30	13.067	8.0276	5.5077	3.9172	3.0515	2.4176	2.0211	1.7096	1.5174	1.2992	1.1765
MR	ARL	82.732	18.157	14.101	11.931	10.247	9.1981	8.2883	7.6439	7.0199	6.6075	6.1599	5.8605
	SDRL	72.299	9.4484	6.0986	4.4767	3.4233	2.7943	2.2216	1.9336	1.6438	1.4571	1.2749	1.1743
TM	ARL	213.82	19.218	14.342	11.950	10.032	8.921	7.9669	7.2973	6.6737	6.2663	5.8517	5.5397
	SDRL	197.50	9.5794	5.9439	4.3042	3.0152	2.4468	1.948	1.635	1.373	1.2218	1.0828	0.974
HL	ARL	192.83	18.635	13.988	11.743	9.8959	8.8148	7.9015	7.2301	6.6191	6.2306	5.825	5.5099
	SDRL	180.19	9.0194	5.5764	4.1339	2.9312	2.3938	1.8944	1.5928	1.3321	1.2001	1.0613	0.9608

Table 4.6: ARL and SDRL Values for the MEDC chart Based on Different estimators Under LCNE with $k_1 = 0.125,\ k_2 = 0.5, \lambda = 0.5$ and different H_1&H_2 for a specified $ARL_0 = 370$

		δ											
$\widehat{\theta}$	Values	0	0.25	0.33	0.4	0.48	0.55	0.63	0.7	0.78	0.85	0.93	1
$\bar{x}$	ARL	131.93	16.027	11.549	9.4114	7.8103	6.7751	5.9968	5.3879	4.9113	4.5644	4.2472	3.9795
	SDRL	124.53	9.5044	5.5904	4.0469	2.9977	2.3462	1.881	1.5427	1.2801	1.1312	0.9994	0.8792
$\tilde{x}$	ARL	273.39	23.720	16.244	12.600	12.600	8.529	7.377	6.6092	5.9157	5.442	5.0103	4.6835
	SDRL	260.83	15.974	9.4635	6.4635	6.4223	3.424	2.6553	2.1934	1.7994	1.5359	1.3126	1.1626
MR	ARL	75.615	16.819	12.692	10.440	8.6822	7.559	6.659	5.9959	5.4721	5.0561	4.7132	4.4327
	SDRL	68.965	10.665	6.9721	5.0664	3.7881	2.9728	2.3737	1.9892	1.6606	1.437	1.2768	1.1348
TM	ARL	214.21	18.488	13.154	10.356	8.4323	7.2705	6.0307	5.6978	5.1534	4.782	4.4267	4.1373
	SDRL	204.57	11.387	6.8726	4.7341	3.3499	2.6397	2.022	1.3998	1.3998	1.2207	1.0597	0.9427
HL	ARL	196.96	17.92	12.758	10.246	8.2887	7.2355	6.2815	5.6972	5.1369	4.7762	4.4128	4.1458
	SDRL	187.67	10.714	6.464	4.5638	3.2277	2.4942	1.9789	1.649	1.3616	1.1751	1.032	0.9094

Table 4.7: ARL and SDRL Values for the MEDC chart Based on Different estimators Under LCNE with $k_1 = 0.125,\ k_2 = 0.5, \lambda = 0.75$ and different H_1&H_2 for a specified $ARL_0 = 370$

		δ											
$\widehat{\theta}$	Values	0	0.25	0.33	0.4	0.48	0.55	0.63	0.7	0.78	0.85	0.93	1
$\bar{x}$	ARL	127.38	16.032	11.279	8.8931	7.1474	6.144	5.2614	4.7021	4.2318	3.8789	3.5611	3.3204
	SDRL	118.40	10.681	6.5631	4.5813	3.2548	2.601	1.9942	1.6134	1.4024	1.1856	1.0227	0.9181
$\tilde{x}$	ARL	254.52	24.891	16.495	12.413	9.6508	7.8909	6.7249	5.8767	5.1643	4.6767	4.2769	3.9692
	SDRL	245.98	17.502	10.640	7.3734	5.184	3.7532	2.9227	2.3743	1.8904	1.6077	1.3954	1.2301
MR	ARL	67.630	16.571	12.113	9.7813	8.0722	6.9168	5.9393	5.3216	4.7552	4.364	3.9852	3.7143
	SDRL	62.485	11.722	7.6147	5.4908	4.1588	3.2751	2.579	2.162	1.753	1.545	1.3135	1.1653
TM	ARL	202.12	18.916	12.879	9.8679	7.7831	6.6105	5.6184	5.0234	4.4208	4.0588	3.6764	3.4428
	SDRL	188.51	12.972	7.9826	5.2652	3.6994	2.8406	2.205	1.8199	1.4711	1.263	1.0946	0.9636
HL	ARL	189.26	18.210	12.499	9.7664	7.6602	6.5255	5.6156	4.9527	4.4403	4.0524	3.7172	3.4673
	SDRL	178.32	12.139	7.368	5.088	3.5104	2.7052	2.148	1.7597	1.4393	1.2425	1.074	0.9525

4.4.3 Variance – Contaminated Normal Environment (VCNE)

A random variable say X, is said to be drawn from a VCNE, if it has $(\varphi)100\%$ observations from $N(\mu_0, \sigma_0^2)$ and $(1-\varphi)100\%$ observations from $N(\mu_0, \tau\sigma_0^2)$ where $-\infty < \tau < \infty$. Similarly, we generated the random variables for simulating the MEDC schemes from 5% and 1% VCNE, with $\tau = 9$, for different values of λ. The outcomes of the simulation results were not different rather better than the LCNE case. Here, the least ARL_0 is greater than 80% of the expected 370 value for the in-control ARL. The hierarchy of the estimators' performance is still maintained in this environment both in the in and out of control ARL's. The trimean, Hodges-Lehmann and median estimators (in hierarchy) are more robust than the mean and midrange estimators.

Table 4.8: ARL and SDRL Values for the MEDC chart Based on Different estimators Under VCNE with $k_1 = 0.125,\ k_2 = 0.5, \lambda = 0.05$ and different $H_1 \& H_2$ for a specified $ARL_0 = 370$

		δ											
$\hat{\theta}$	Values	0	0.25	0.33	0.4	0.48	0.55	0.63	0.7	0.78	0.85	0.93	1
$\bar{x}$	ARL	339.40	24.667	19.658	16.758	14.519	13.074	11.826	10.952	10.11	9.5261	8.9014	8.4446
	SDRL	320.72	8.7447	5.5657	4.1503	3.1479	2.6121	2.1364	1.823	1.5788	1.4331	1.2676	1.1448
$\tilde{x}$	ARL	359.68	28.334	22.420	18.979	16.340	14.679	13.291	12.260	11.296	10.620	9.9345	9.4138
	SDRL	342.98	11.292	7.3163	5.2487	3.9441	3.2403	2.6767	2.2829	1.9598	1.7756	1.558	1.4062
MR	ARL	311.92	28.360	23.396	19.115	16.536	14.857	13.411	12.384	11.397	10.727	10.061	9.5131
	SDRL	293.85	11.417	7.2101	5.5627	4.0915	3.3907	2.7935	2.3958	2.0255	1.82	1.6004	1.4531
TM	ARL	368.85	25.666	20.249	17.342	15.021	13.560	12.175	11.320	10.427	9.7934	9.1596	8.6956
	SDRL	348.84	9.1519	5.8679	4.3964	3.31	2.7343	2.2458	1.9352	1.6606	1.4609	1.292	1.1828
HL	ARL	357.54	25.416	20.043	17.114	14.923	13.410	12.116	11.207	10.367	9.744	9.1072	8.6553
	SDRL	333.44	8.8318	5.6761	4.2372	3.233	2.6435	2.1819	1.909	1.6398	1.4573	1.2796	1.1604

Table 4.9: ARL and SDRL Values for the MEDC chart Based on Different estimators Under VCNE with $k_1 = 0.125,\ k_2 = 0.5, \lambda = 0.25$ and different H_1&H_2 for a specified $ARL_0 = 370$

		δ											
$\widehat{\theta}$	Values	0	0.25	0.33	0.4	0.48	0.55	0.63	0.7	0.78	0.85	0.93	1
$\bar{x}$	ARL	320.16	20.314	14.824	12.126	10.099	8.9419	7.9392	7.2973	6.7001	6.2356	5.7967	5.4972
	SDRL	307.96	10.462	6.1442	4.3288	3.0451	2.4369	1.9125	1.6262	1.3747	1.2016	1.0521	0.9596
$\tilde{x}$	ARL	363.25	25.195	17.770	14.301	11.880	10.377	9.1604	8.3286	7.5639	7.0511	6.5504	6.1897
	SDRL	351.23	14.622	8.3449	5.7833	4.0771	3.224	2.5291	2.084	1.7363	1.5304	1.3313	1.1948
MR	ARL	259.54	23.970	17.405	13.935	11.608	10.228	9.0546	8.2403	7.5276	7.0014	6.5459	6.1786
	SDRL	243.82	13.751	8.403	5.58	4.0638	3.2183	2.5472	2.1693	1.0897	1.5736	1.3739	1.2434
	ARL	362.47	21.404	15.457	12.651	10.487	9.2637	8.203	7.5058	6.8259	6.4073	5.9679	5.6357
	SDRL	342.76	11.314	6.7192	4.7225	3.2659	2.5954	2.0395	1.7119	1.4231	1.2623	1.1098	0.9983
HL	ARL	341.43	20.935	15.153	12.367	10.345	9.1303	8.0925	7.4308	6.7959	6.3549	5.9289	5.6187
	SDRL	330.48	10.989	6.1922	4.3821	3.1299	2.4803	1.9596	1.6519	1.4053	1.2365	1.0649	0.9709

Table 4.10: ARL and SDRL Values for the MEDC chart Based on Different estimators Under VCNE with $k_1 = 0.125,\ k_2 = 0.5, \lambda = 0.25$ and different H_1&H_2 for a specified $ARL_0 = 370$

		δ											
$\widehat{\theta}$	Values	0	0.25	0.33	0.4	0.48	0.55	0.63	0.7	0.78	0.85	0.93	1
$\bar{x}$	ARL	303.60	19.625	13.348	10.536	8.5411	7.2643	6.3499	5.6781	5.1703	4.7287	4.374	4.1055
	SDRL	297.10	12.214	7.1052	4.7662	3.415	2.5558	2.0234	1.6669	1.4025	1.1768	1.0345	0.9196
$\tilde{x}$	ARL	358.67	25.715	17.174	13.160	10.514	8.8422	7.5685	6.7336	5.9904	5.5259	5.0717	4.7456
	SDRL	349.69	17.964	10.023	6.7876	4.7717	3.5748	2.7345	2.2888	1.8432	1.5797	1.3212	1.1848
MR	ARL	229.28	23.285	16.119	12.583	10.068	8.552	7.3906	6.605	5.8993	5.4378	5.0014	4.6679
	SDRL	218.78	15.689	9.3596	6.3868	4.5438	3.4738	2.696	2.2156	1.8018	1.5633	1.3335	1.19
	ARL	356.67	21.212	14.308	11.199	8.8874	7.6336	6.5769	5.8981	5.2657	4.9113	4.5226	4.2341
	SDRL	345.29	13.825	7.8759	5.4249	3.6154	2.834	2.1665	1.7837	1.4396	1.2619	1.0962	0.9722
HL	ARL	351.73	22.868	15.339	11.783	9.4207	8.0807	6.9471	6.2213	5.5987	5.1903	4.7552	4.4498
	SDRL	345.91	14.501	8.1439	5.3555	3.6595	2.8693	2.1835	1.8036	1.4965	1.3019	1.0859	0.9809

Table 4.11: ARL and SDRL Values for the MEDC chart Based on Different estimators Under VCNE with $k_1 = 0.125,\ k_2 = 0.5, \lambda = 0.75$ and different H_1&H_2 for a specified $ARL_0 = 370$

		δ											
$\widehat{\theta}$	Values	0	0.25	0.33	0.4	0.48	0.55	0.63	0.7	0.78	0.85	0.93	1
$\bar{x}$	ARL	289.76	20.010	13.251	10.089	7.8081	6.6555	5.6345	5.0124	4.4086	4.0346	3.6867	3.4276
	SDRL	276.13	13.745	7.9303	5.4445	3.6418	2.8353	2.1643	1.7897	1.4492	1.2321	1.0731	0.9437
$\tilde{x}$	ARL	340.46	27.229	17.622	13.044	10.024	8.1992	6.8737	6.0247	5.2992	4.7898	4.3306	4.0345
	SDRL	332.39	19.476	11.686	7.877	5.3606	4.0018	3.0832	2.4348	1.9755	1.7057	1.4241	1.2576
MR	ARL	208.73	24.037	16.033	12.200	9.4371	7.8995	6.6401	5.8558	5.1664	4.7333	4.2544	3.9594
	SDRL	205.14	17.674	10.594	7.1434	4.968	3.74	2.9132	2.3641	1.9167	1.6496	1.3763	1.2245
TM	ARL	329.35	22.100	14.324	10.651	8.3228	6.9218	5.8689	5.1781	4.5922	4.1559	3.7839	3.509
	SDRL	319.83	15.569	9.0028	5.9133	3.9524	3.0629	2.3216	1.9021	1.5608	1.3239	1.1342	0.9815
HL	ARL	336.63	21.276	13.922	10.438	8.2275	6.9625	5.8332	5.1375	4.5839	4.164	3.7998	3.5417
	SDRL	330.32	14.779	8.4063	5.5027	3.8741	3.0169	2.2404	1.8133	1.5078	1.3066	1.1113	0.97

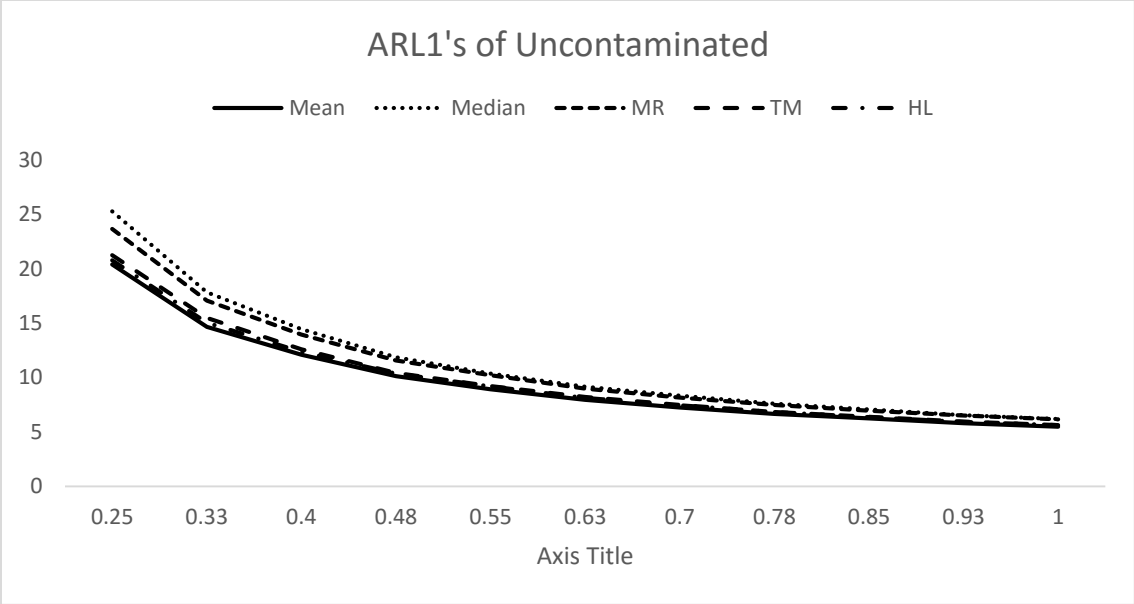

Figure 4.0: The ARL_1's curve of the MEDC uncontaminated environment

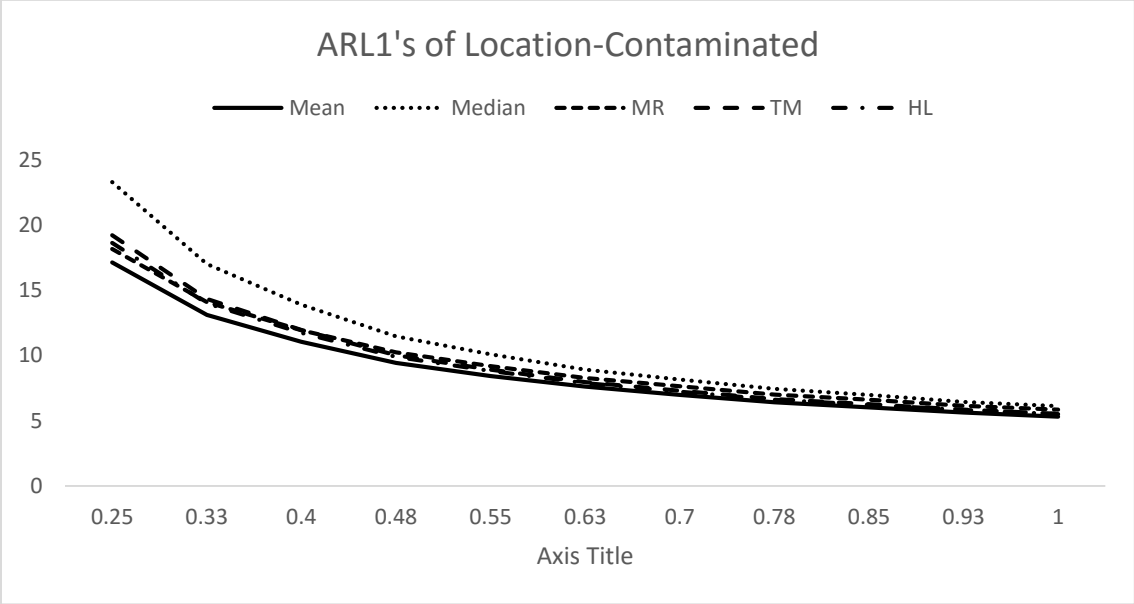

Figure 4.1: The ARL_1's curve of the MEDC for LCNE

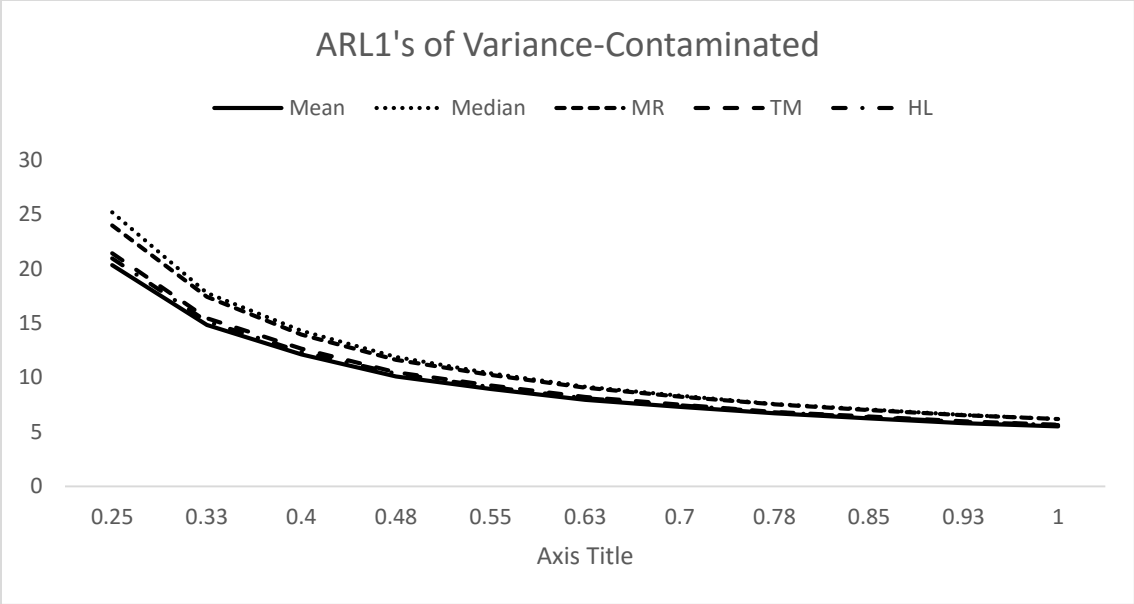

Figure 4.2: The ARL_1's curve of the MEDC VCNE

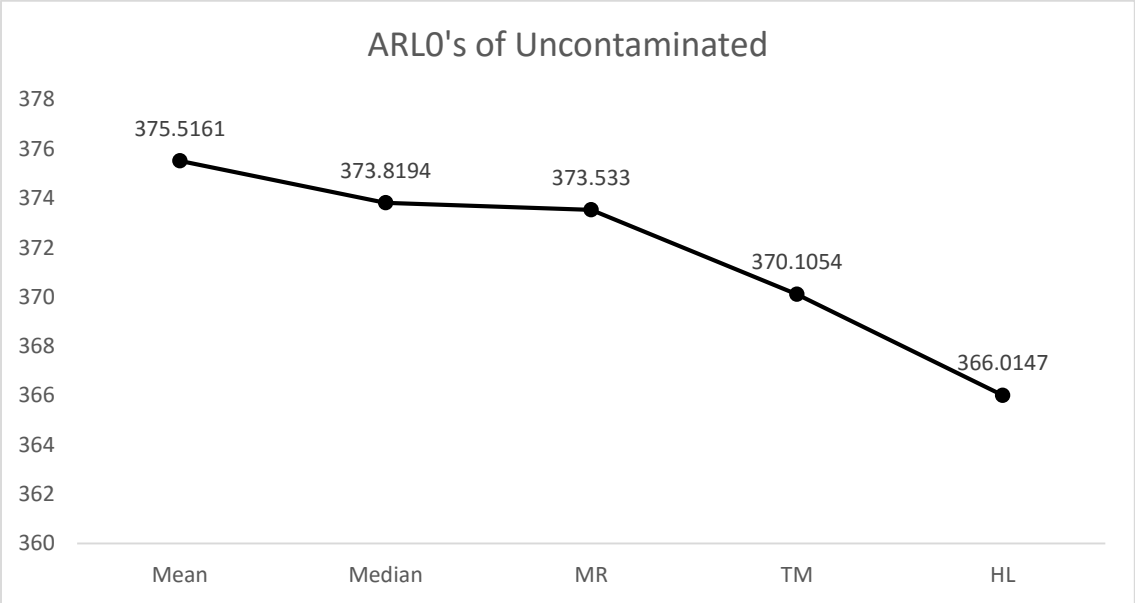

Figure 4.3: The ARL_0's curve of the MEDC for uncontaminated environment

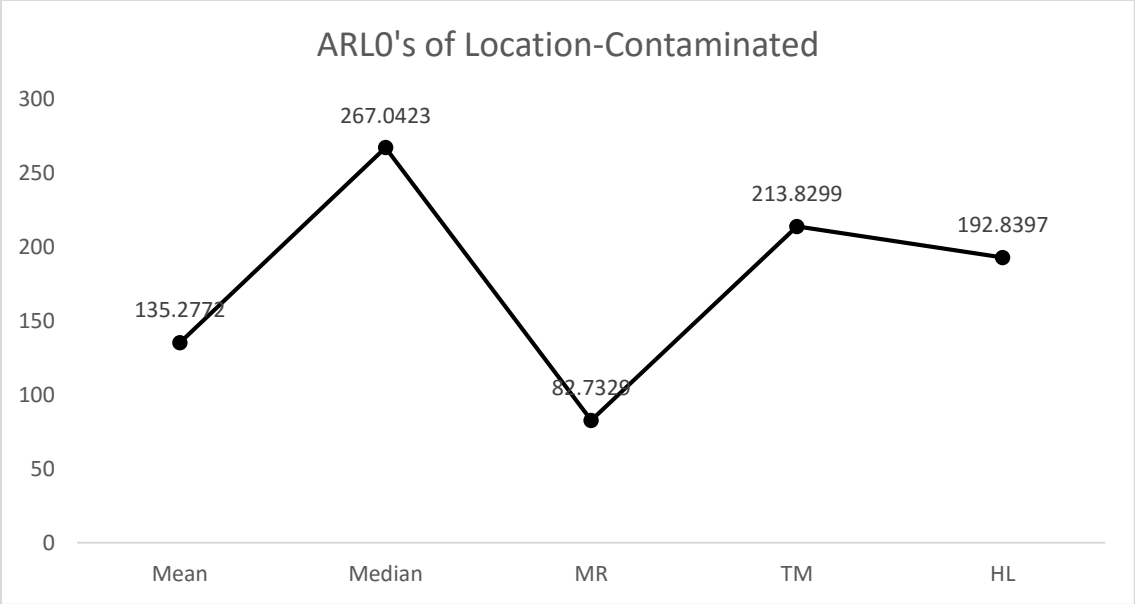

Figure 4.4: The ARL_0′s curve of the MEDC for LCNE

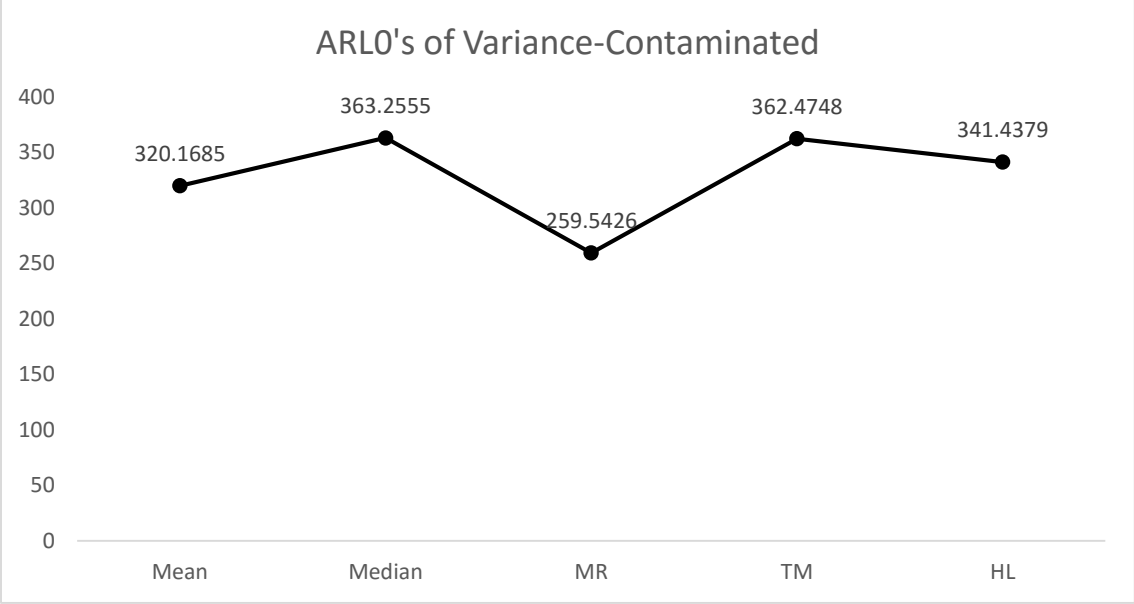

Figure 4.5: The ARL_0′s curve of the MEDC for VCNE

4.5 APPLICATION OF MEDC SCHEME WITH A REAL-LIFE DATA

In this section, we present a description of the real-life dataset we applied our new proposed MEDC scheme on and how this data set was applied with the scheme, to illustrate the robustness of each of the five estimators; mean, median, midrange, trimean and Hodges-Lehmann.

Chloride is the active component of the catalyst used in the Naphtha Reforming Process at petroleum refineries; where naphtha quality is upgraded from low octane number naphtha to high octane naphtha upon passing this low naphtha feed via reactors filled with catalyst. The high octane number naphtha is used for making high octane number gasoline.

Catalysts having very high concentration of chloride will cause severe cracking to the feed (i.e. the naphtha) and catalyst with very low centration of chloride will make the catalyst inactive. So the chloride in catalyst concentration limit have to be balanced and controlled during the chemical reforming process. Therefore, the variable of interest to be monitored and measured in the process is the concentration of chloride.

The catalyst is very expensive and shutting down the naphtha reforming chemical plant due to uncontrolled chloride concentration impacts negatively the chemical process plant economics. So regular catalyst sampling and testing for the chloride content is a common practice at petroleum refineries. Typically, X-Ray Fluorescence Spectrometry (XRFS) is the technique used at petroleum refineries laboratories for testing the naphtha reforming catalyst samples for their chloride content. This lab testing procedure is based on grinding small portion of catalysts to be a fine powder, pack them and then introduce them in the XRFS for measuring the chloride concentration level. A Shewhart chart type is used for monitoring the XRFS chloride analyzer for potential process drifts.

As we are acquainted that the Shewhart type of chart is adequate for monitoring large shifts, using its type of chart may not be appropriate in detecting slight drifts in the level of concentration of chloride in this process.

However, we employ our MEDC scheme which is more sensitive to small shifts. The dataset which is the concentration level of chloride in a catalyst sample measured in wt% has 30 sample points each of sample size $n = 5$. We calculated the five aforementioned estimators from the original dataset without contaminations and used them to construct their respective MEDC statistics. The statistics are then plotted against their limits. We introduced some contaminations both location and variance to evaluate their robustness based on the five estimators.

Under the original dataset (uncontaminated), Figures 4.7-4.11 depict the behavior of each of the five estimators when in used in constructing the MEDC scheme with the dataset. A control chart is said to be out-of-control, if its statistics get beyond their limits such as C_1 crossing H_1 and it is said to be in-control if otherwise. Despite the absence of contamination, the mean and midrange are in out-of-control state as shown in Figures 4.7 and 4.9 respectively. While other estimators median, trimean and Hodges-Lehmann are doing fine with the scheme with raising any false alarms.

On the other hand, while introducing a 5% location and variance contaminations to the original dataset, in order to study how these estimators could withstand disturbances, we found that the three estimators; median, trimean and Hodges-Lehmann still maintain their efficiencies and could contain the disturbances. This implies that these three estimators are robust, unlike the mean and median which could not maintain in the original dataset let

alone the contaminated data set. The Figures 4.12-4.16 display the control charts of the respective five estimators with the MEDC schemes for the location-contaminated dataset. While the remainder of the figures are for the variance-contaminated dataset. (i.e. Figures 4.17-4.21).

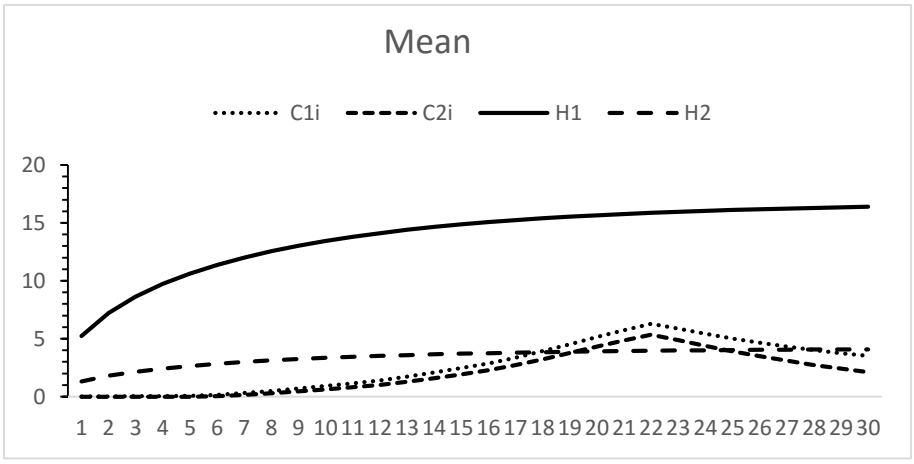

Figure 4.6: Control Chart of mean MEDC statistics for uncontaminated dataset

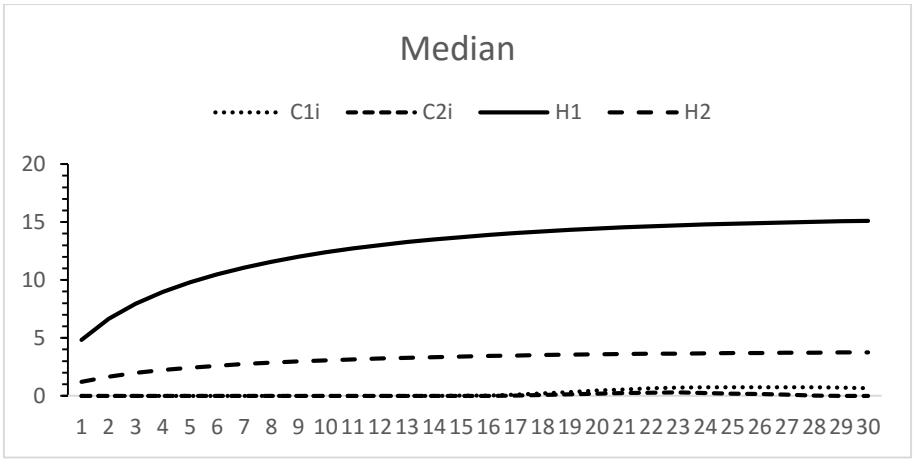

Figure 4.7: Control Chart of median MEDC statistics for uncontaminated dataset

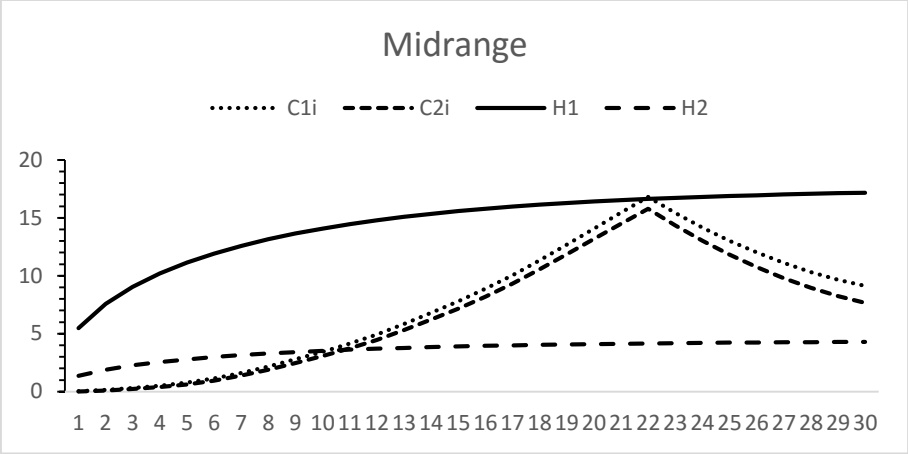

Figure 4.8: Control Chart of midrange MEDC statistics for uncontaminated dataset

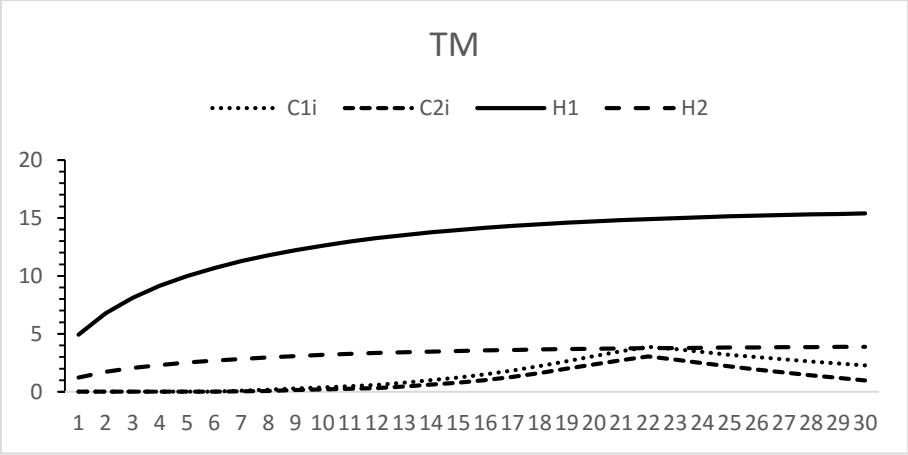

Figure 4.9: Control Chart of TM MEDC statistics for uncontaminated dataset

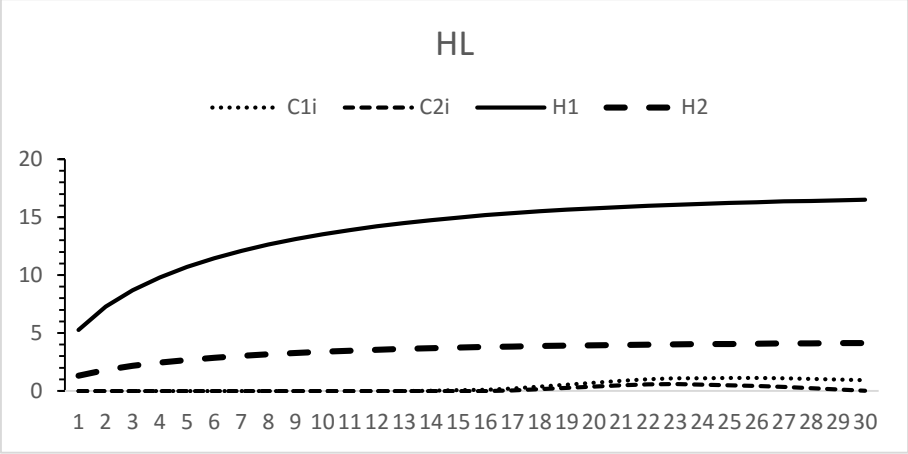

Figure 4.10: Control Chart of HL MEDC statistics for uncontaminated dataset

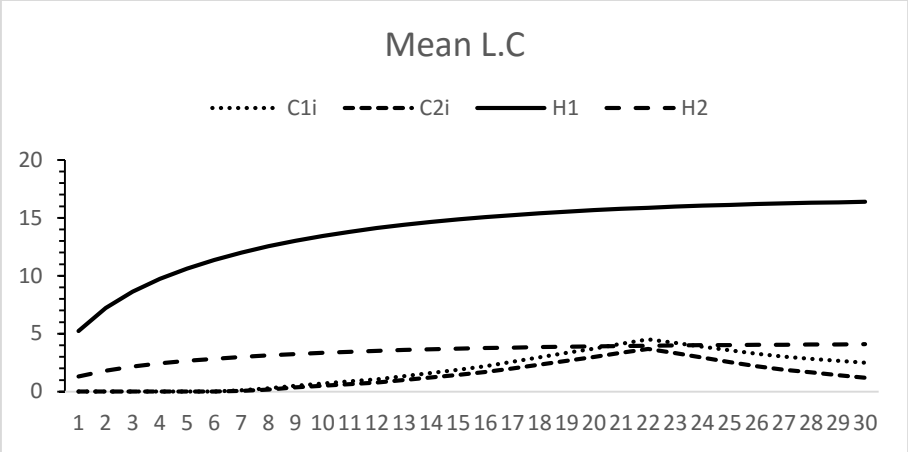

Figure 4.11: Control Chart of mean MEDC statistics for LCNE

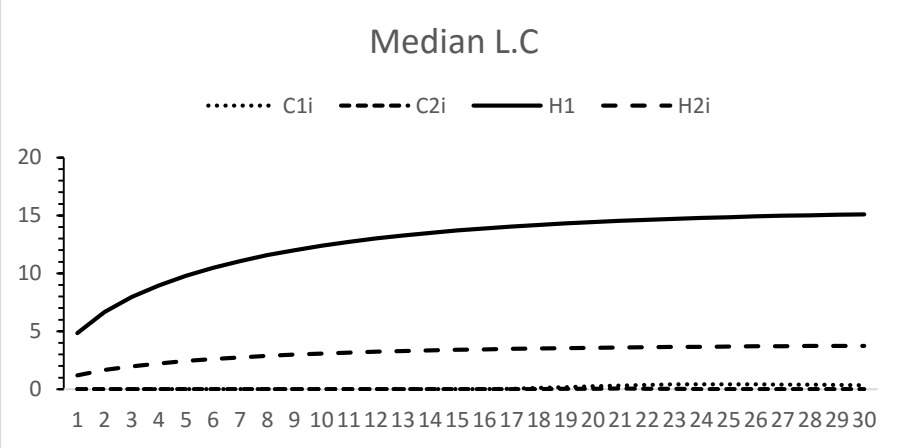

Figure 4.12: Control Chart of median MEDC statistics for LCNE

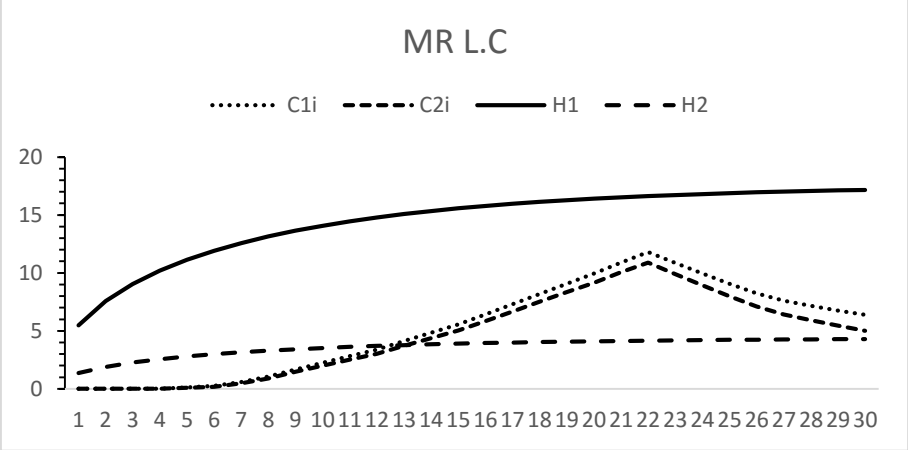

Figure 4.13: Control Chart of midrange MEDC statistics for LCNE

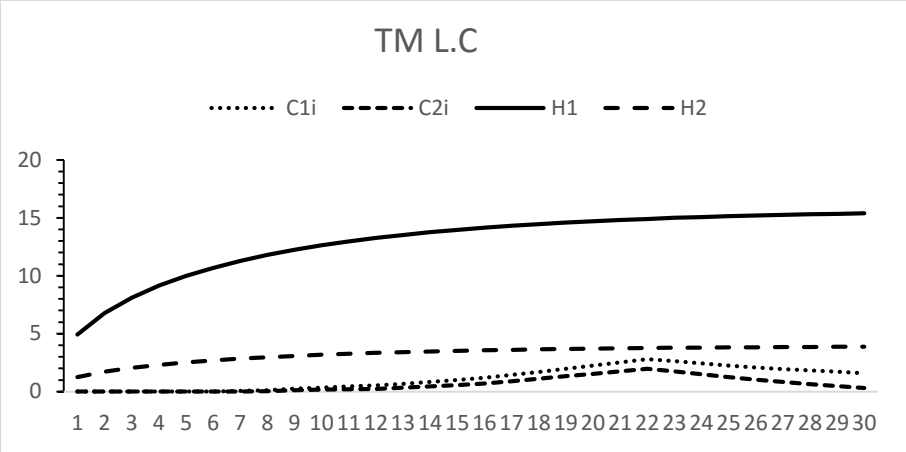

Figure 4.14: Control Chart of trimean MEDC statistics for LCNE

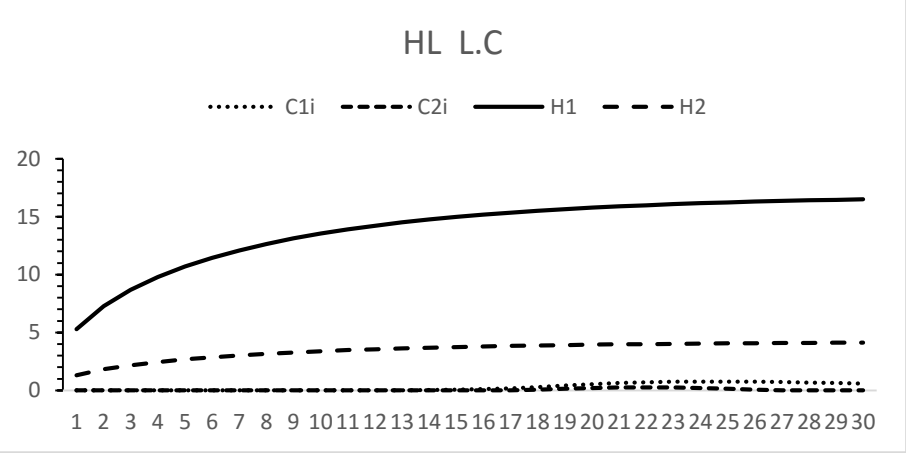

Figure 4.15: Control Chart of HL MEDC statistics for LCNE

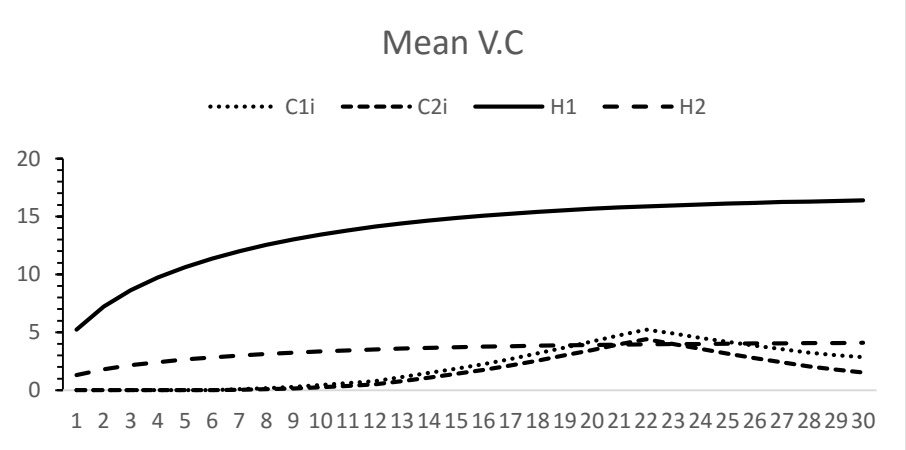

Figure 4.16: Control Chart of mean MEDC statistics for VCNE

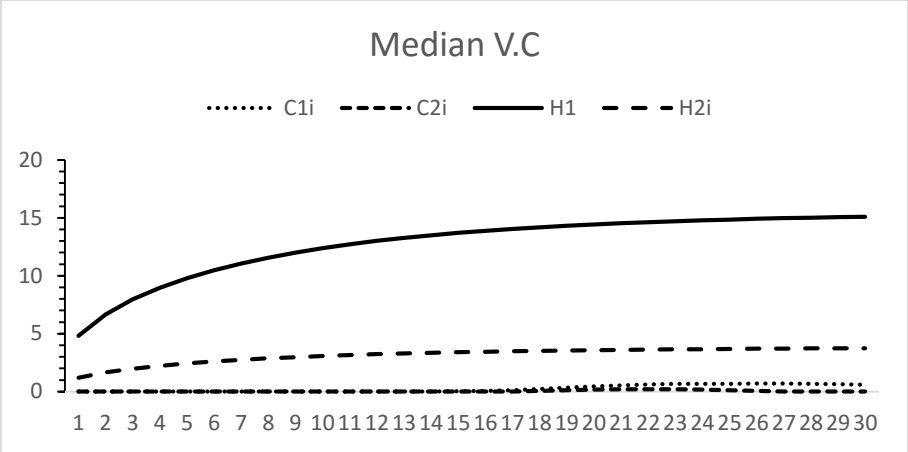

Figure 4.17: Control Chart of median MEDC statistics for VCNE

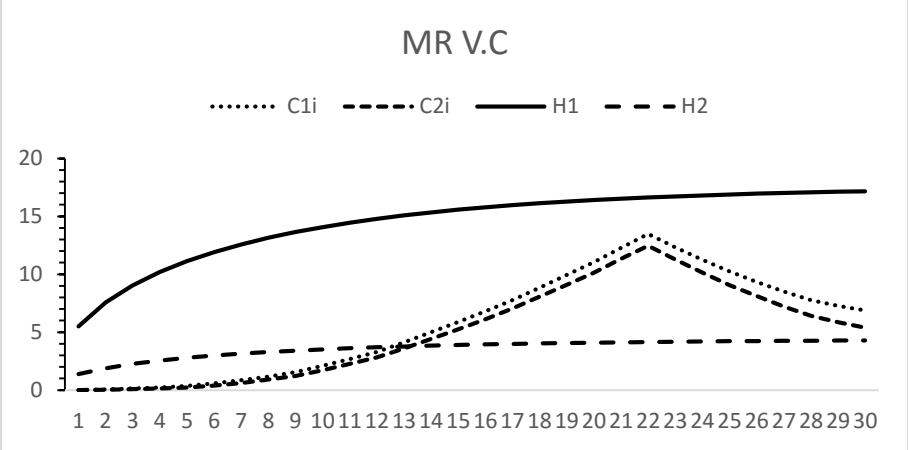

Figure 4.18: Control Chart of midrange MEDC statistics for VCNE

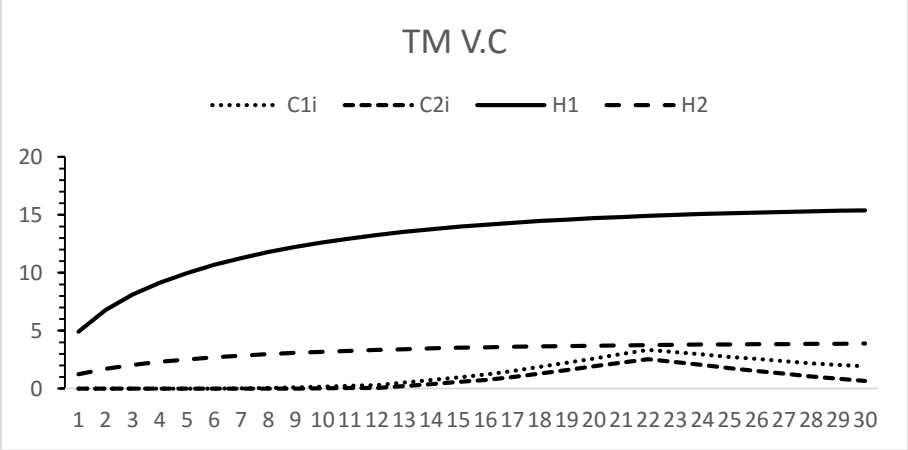

Figure 4.19: Control Chart of trimean MEDC statistics for VCNE

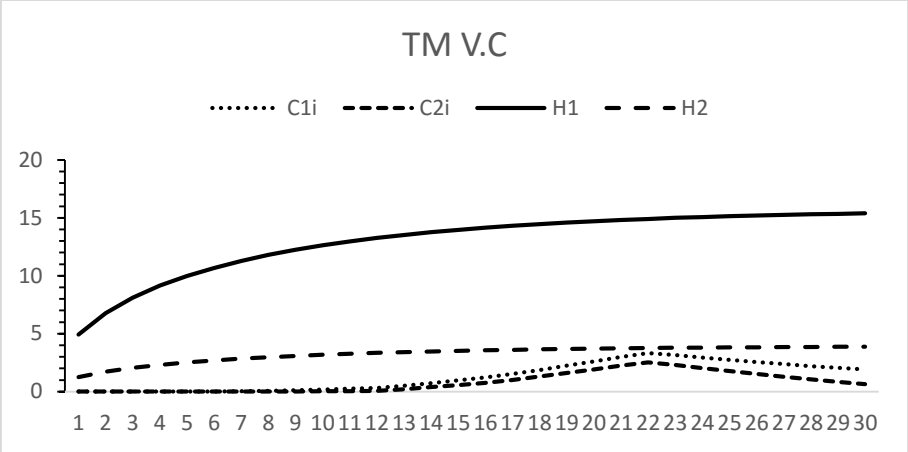

Figure 4.20: Control Chart of HL MEDC statistics for VCNE

4.6 COMPARISON OF MCDE SCHEME WITH COUNTER PARTS

In this section, we make a brief comparison between the Mixed Dual-CUSUM EWMA schemes and DCUSUM and DEWMA charts respectively. To avoid repeating Tables, we shall reference the Tables by their numbers for the purpose of comparison. A relatively balanced platform of comparison could be reached by selecting close parameters of all the schemes. MCDE has both structures of EWMA and CUSUM within it, so this make it possible, since we carried out all the schemes for the same $ARL_0 = 370.$. Another advantage is that all the estimators retained the position in terms of performance in all the schemes. Therefore, the comparison goes straightforward.

Under the uncontaminated normal environment, we close our eyes on the ARL_0 values, and emphasize on the out-of-control values. Recall Tables 3.1 and 4.1 for Double EWMA and MEDC charts respectively, it's obvious the DEWMA chart has lower ARL_1 values and decays faster as the shift increases compared to the MEDC chart. Comparing MEDC chart with DCUSUM chart, using Tables 4.1 and 2.1 respectively, the DCUSM also outperforms the MEDC scheme.

For the contaminated environments, both location and variance contamination of 1%, according to Tables 3.2 and 4.5 for DEWMA versus MCDE, with respect to location contamination, the former crushes the latter both ARL_0 and ARL_1 values. On the other hand, MCDE trounces DCUSUM chart while comparing Tables 4.5 and 2.2 respectively.

Surprisingly, the synopsis changes under the variance-contaminated environment, as MCDE scheme thrashes the DEWMA chart let alone the DCUSUM chart for both ARL_0 and ARL_1 values. Tables 4.9, 3.4 and 2.4 justify this claim respectively.

4.7 SUMMARY AND CONCLUSION

In this chapter, we proposed a mixture of the DCUSUM and EWMA structures, making DCUSUM a dominant part of the scheme. We explored the scheme with the five estimators and investigate their robustness under contaminated and uncontaminated normal processes. We made a brief comparison of the scheme with its counterparts, DCUSUM and DEWMA respectively. With the EWMA structure embedded, the smoothing factor is also playing a significant role to how fast, the scheme could detect small shifts if any.

Findings show that the mean estimator is preferred to others under the uncontaminated environments, but it is not robust against contamination. Under the location and variance contamination, the trimean estimator is out performs others, next to it in performance is the Hodges-Lehmann estimator. Despite the largest values of the median estimator for the in-control ARL, its weak performance in the out-of-control deprived it from been the best in the contaminated environments. Therefore, trimean estimator is the best overall.

It should be noted that the MEDC scheme is very robust against contamination. It has almost the same run length values with the uncontaminated or even better. Also its point where the MEDC is taking lead when compared with the DCUSUM and DEWMA schemes.

CHAPTER 5

GENERAL CONCLUSION

In this study, we have carefully investigated the robustness and efficiency of some memory type control charts in dual forms. They are DCUSUM and DEWMA charts for location monitoring. These two charts were usually constructed by mean estimator for subgroup observations, whereas, our study proves that the mean estimator performs woefully in presence of little or no disturbances in the process under consideration. In order to increase the efficiency of these scheme, we proposed some location estimators; mean, median, midrange, trimean and Hodges-Lehmann. The mean estimator was kept in the study as a yardstick for performance evaluation. These estimators were therefore used to construct the DCUSUM and DEWMA schemes, both in a normal environment (uncontaminated) and contaminated environments.

With the aid of Monte Carlo simulation, we derived the ARL as the performance measure. The simulation results showed that the mean and midrange are less robust as compared to the remaining three estimators. Amongst the median, trimean and Hodges-Lehmann, the trimean estimator is preferred. Though the median estimator is posing to have good performance for the in-control ARL, but it's less efficient in quickly detecting small drifts in processes under control. This deficiency in efficiency, made median lost the position to trimean and Hodges-Lehmann.

The application of these schemes on a real-life data set from regional sector, buttressed the claims further, where the weakness of the mean and midrange estimators were depicted with control chart figures.

In addendum, we proposed a new control scheme, which was an assortment of a dual-memory chart and a classical memory chart, named Mixed EWMA-DUAL-CUSUM (MEDC) for location monitoring. The scheme was also explored with the five estimators to evaluate its robustness and efficiency in the presence of disturbances (contamination) and violation of normality assumptions. The same conclusion also hold with respect to the five estimators. We applied the scheme on a real-life data set from petroleum refineries, in order to measure/monitor the concentration level of chloride in a catalyst.

We had a comparative study of these three schemes, DCUSUM, DEWMA and MEDC, and findings showed that the new proposed scheme is preferred. More so it is more robust against variance contaminations.

References

[1] D. Montgomery, *Introduction to statistical quality control*. 2009.

[2] W. A. Shewhart, "Economic control of quality of manufactured product,New York," *MacMillan, London*, 1931.

[3] E. S. Page, "Continuous Inspection Schemes," *Biometrika*, vol. 41, 1954.

[4] S. W. Roberts, "Control Chart Tests Based on Geometric Moving Averages Technomearics," *Technomearics*, 1959.

[5] J. M. Lucas, "Coombined Shewhart-CUSUM quality control schemes," *J. Qual. Technol.*, 1982.

[6] M. S. Lucas, James M & Saccucci, "Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements.," *Technometrics*, 1990.

[7] R. B. Lucas, James M. & Crosier, "Fast Initial Response for CUSUM Quality-Control Scheme.," *Technometrics*, vol. 24, pp. 199–205, 1982.

[8] N. Abbas, M. Riaz, and R. J. M. M. Does, "Mixed Exponentially Weighted Moving Average-Cumulative Sum Charts for Process Monitoring," *Qual. Reliab. Eng. Int.*, vol. 29, no. 3, pp. 345–356, 2013.

[9] R. Domangue and S. C. Patch, "Some Omnibus Exponentially

Weighted Moving Average Statistical Process Monitoring Schemes," *Technometrics*, Mar. 2012.

[10] G. Capizzi, G. and Masarotto, "An Adaptive Exponentially Weighted Moving Average Control Chart," *Technometrics*, vol. 45, no. 3, pp. 199–207, 2003.

[11] & Borror, C. M., Montgomery, D. C. and G. C. Runger, "Robustness of the EWMA control chart to non-normality," *J. Qual. Technol.*, vol. 10, pp. 139–149, 1999.

[12] Y. Zhao, F. Tsung, and Z. Wang, "Dual CUSUM control schemes for detecting a range of mean shifts," *IIE Trans.*, vol. 37, no. 11, pp. 1047–1057, Nov. 2005.

[13] G. Lorden, "Procedures for Reacting to a Change in Distribution.," *Ann. Math. Stat.*, vol. 42, no. 6, pp. 1897–1908, 1971.

[14] L. G. Tatum, "Robust Estimation of the Process Standard Deviation for Control Charts," *Technometrics*, vol. 39, pp. 127–141, 1997.

[15] B. A. Moustafa, A. & Mokhtar, "New Robust Statistical Process Control Chart for Location," *Qual. Eng.*, vol. 12, no. 2, pp. 149–159, 1999.

[16] M. Riaz, N. Abbas, and R. J. M. M. Does, "Improving the performance of CUSUM charts," *Qual. Reliab. Eng. Int.*, vol. 27, no. 4, pp. 415–424, Jun. 2011.

[17] R. J. M. M. Schoonhoven, M., Nazir, H.Z., Riaz, M. and Does,

"Robust Location Estimators for the $\bar{X}$ control chart.," *J. Qual. Technol.*, vol. 43, pp. 363–379, 2011.

[18] Y. . Yang, L., Pai, S., and Wang, "A Novel CUSUM Median Control Chart," in *Proceedings of International Multiconference of Engineers and Computer Scientists, Hong Kong.*, 2010.

[19] H. Z. Nazir, M. Riaz, R. J. M. M. Does, and N. Abbas, "Robust CUSUM Control Charting," *Qual. Eng.*, vol. 25, no. 3, pp. 211–224, Jul. 2013.

[20] R. J. M. . Nazir, H. Z., Abbas, N., Riaz, M. and Does, "A Comparative Study of Memory-Type Control Charts under Normal, Contaminated and Non-Normal Environments," *Qual. Technol. Qual. Manag.*, 2014.

[21] E. L. Hodges, J. L. Jr. and Lehmann, "Estimates of location based on rank tests," *Ann. Math. Stat.*, vol. 34, pp. 598–611, 1963.

[22] C. W. Jones-Farmer, L. A., Jordan, V., Champ, "Distribution-free phase I charts for subgroup location," *J. Qual. Technol.*, vol. 41, no. 3, pp. 304–316, 2009.

[23] A. K. Shamma, S. E., & Shamma, "Development and evaluation of control charts using double exponentially weighted moving averages," *Int. J. Qual. Reliab. Manag.*, vol. 9, no. 6, pp. 18–25, 1992.

[24] L. Y. Zhang, "EWMA control charts and extended EWMA control charts," *Unpubl. Dr. Diss. Univ. Regina, Saskatchewan, Canada*, 2002.

[25] M. a. Mahmoud and W. H. Woodall, “An Evaluation of the Double Exponentially Weighted Moving Average Control Chart,” *Commun. Stat. - Simul. Comput.*, vol. 39, no. 5, pp. 933–949, 2010.

[26] S. S. Alkahtani, “Robustness of DEWMA versus EWMA Control Charts to Non-Normal Processes,” vol. 12, no. 1, pp. 148–163, 2013.

[27] R. G. Brown, “Smoothing, Forecasting and Prediction,” in *Englewood Cliffs, NJ: Prentice*, 1962.

[28] G. Zhang, L., & Chen, “An extended EWMA mean chart,” *Qual. Technol. Quant. Manag.*, vol. 2, pp. 39–52, 2005.

[29] S. H. Steiner, “EWMA Control Charts with Time Varying Control Limits and Fast Initial Response,” *J. Qual. Technol.*, vol. 31, no. 1, pp. 75–86, 1999.

[30] E. Yashchin, “Weighted Cumulative Sum Technique,” *Technometrics*, vol. 31, pp. 321–338, 1989.

[31] W. Jiang, L. Shu, and D. W. Apley, “Adaptive CUSUM procedures with EWMA-based shift estimators,” *IIE Trans.*, vol. 40, no. 10, pp. 992–1003, 2008.

[32] R. J. M. . Abbas, N., Riaz, M., and Does, “Enhancing the Performance of EWMA Charts,” *Qual. Reliab. Eng. Int.*, vol. 27, no. 6, pp. 821–833, 2011.

[33] B. Zaman, M. Riaz, N. Abbas, and R. J. M. M. Does, “Mixed Cumulative Sum-Exponentially Weighted Moving Average

Control Charts: An Efficient Way of Monitoring Process Location," *Qual. Reliab. Eng. Int.*, p. n/a–n/a, Jun. 2014.

Vitae

Name : Raji Ishaq Adeyanju

Nationality : Nigerian

Date of Birth : 23/06/1988

Email : bunraj0105@gmail.com

Address : 9A, Baale Street Ishagatedo via Isolo, Lagos, Nigeria.

Academic Background :

- M.Sc. Applied Statistics, King Fahd University of Petroleum & Minerals, KFUPM, Dhahran, Saudi Arabia. 2013 -2015
- B.Sc. Statistics, University of Ilorin, Ilorin, Kwara State, Nigeria. 2006-2010.

Submitted Publications Under Peer Review:

- Robust Dual-CUSUM Control Charts for Contaminated Processes. "*Computers & Industrial Engineering",* Raji Ishaq Adeyanju, Mohammad Riaz and Nasir Abbas
- On Designing a Robust Double EWMA Control Chart for Process Monitoring. "*Communications in Statistics-Theory and Methods",* Raji Ishaq Adeyanju, Nasir Abbas and Mohammad Riaz.
- An Application of Mixed EWMA DUAL-CUSUM Scheme in Petroleum Refinery and Health Sector. "*Journal of Applied Statistics",* Raji Ishaq Adeyanju, Nasir Abbas, Mohammad Riaz and Khalid Al-Ghamdi.